Polarised spectroscopy to probe biomacromolecular structure and function

ALISON RODGER

DOCTOR OF SCIENCE

UNIVERSITY OF WARWICK

2007

1. INTRODUCTION

The work presented in this thesis has been published by Alison Rodger (AR) over the last five years. It demonstrates the culmination of twenty five years of research undertaken both independently and in collaboration with colleagues around the world, including thirty two PhD students and many Masters students. The full publication list is given in Appendix 1. Most of the work prior to 2002 was considered for a DSc from Sydney University, so is not included in this submission. The publications for consideration have been grouped into: reviews, innovative papers which establish new lines of work, collaborative work undertaken because external researchers sought expertise unique to Warwick, and other papers. The percentages on the right hand side are an estimate of AR's contribution.

The review papers reflect invitations to write the review given by each title. To ensure a balanced perspective they are usually multiple author products.

The papers which establish new lines of work can be grouped into applications and instrumentation. The latter has been driven by (i) the need to reduce sample volume from the previous cm^3 required to mm^3, (ii) the need to improve sensitivity, and (iii) the need to remove artifacts from spectral data collected. The external collaborations established show how widely recognized this work has been. The application innovations have usually arisen as the result of discussions at meetings and conferences when possibilities of polarized linear dichroism (the difference in absorption of light polarized parallel and perpendicular to an orientation axis) for someone elses project have been explored. New systems studied, which are all leading to a mushrooming of collaborations, include carbon nanotubes, liposomes (artificial cell membrane systems), fibrous proteins, and DNA-protein systems (especially enzymatic ones).

As the titles of the collaborative papers indicate, the applications areas are diverse but they are united by including data and its interpretation that could only have been collected in the unique biophysical spectroscopy facilities which have been established by AR at Warwick. The collaborators come from all round the world. Sometimes the project has proceeded by them visiting to use instruments and sometimes AR or a member of her group undertakes the measurements for them. Her best work comes from collaborating with biologists and chemists who have challenging applications problems which lead to new instrumentation methodological developments. These advances are then implemented in a range of collaborative projects on diverse systems. Most of her students work in inter-university collaborations to the benefit of all concerned.

2. THEME UNDERLYING WORK SUMMARISED IN THE PUBLICATIONS LISTED IN SECTION 3

Over the last few years Alison Rodger's (AR's) laboratory at Warwick has become the national and indeed international hub for innovation based on the technique of flow linear dichroism (*LD*), which is the difference in absorbance of light polarised parallel and perpendicular to an orientation axis. This is what is given in detail in the publications of §3. The key information that *LD* gives is the orientation of sub-units of a system with respect to the orientation axis — which in the case of flow orientation is the long axis of the system. *LD* can be used to deduce kinetic and structural information about a wide range of systems, the only requirements being that they have absorbance spectroscopy and they can be oriented. Flow *LD* requires samples to be long enough to be oriented by shear forces in solution and is ideally suited to DNA and fibrous proteins. More recently we have had significant success in orienting membrane systems as the flow-distortion of liposomes (model membrane systems) creates an orientation axis. In the context of ligand binding, the key attraction of *LD* is that it is selective only for those molecules bound to the long system. Thus, *e.g.*, only elongated polymerase chain reaction products are detected, not the interfering background population of free nucleotides.[17] *LD*, as with other spectroscopic techniques, may have samples in near-native environments.

The key enabling invention for most of the work undertaken in collaboration with external scientists has been AR's invention of a micro-volume Couette *LD* cell, which reduced sample requirements by about 2 orders of magnitude. This work is summarized in references 3 and 7. With the new cell the technique of fluorescence detected linear dichroism which AR had invented was for the first time able to be measured.[7] This invention also improved signal to noise on *LD* data due in large part to the lens system that was required to reduce the size of the light beam to enable it to pass through the micro-volume capillaries. With this cell applications in a range of new areas became immediately possible where they had bee prohibitively expensive previously. The uniqueness of the Warwick instrumentation is indicated by the 18 papers published since 2003 on samples brought to Warwick by external collaborators, most of them new collaborators responding to literature reports of the techniques possibilities.[8–25] AR's own interests are focused on using polarized optical spectroscopy to study the structure and function of biomacromolecules, including DNA-ligand interactions,[R2,R7,27–33] DNA-protein complexes and their chemistry,[2] protein fibres [4,R3,R4] and membrane proteins,[1,R6] soluble proteins and their complexes,[R1,R5] and carbon nanotubes.[5,6] Many of the applications initiated by external collaborators have resulted in innovations in techniques and theory undertaken by AR to the benefit of a wide range of projects. The FtsZ (the bacterial homologue of tubulin) series of papers is most notable in this regard.[9,15,18]

Five papers of particular note were selected for the initial application as they place AR's work in a wider context.

R1: "Circular and linear dichroism of proteins": This paper is a review of the theory of circular dichroism (*CD*) calculations and the work that the authors have undertaken recently to transform the *CD* formalism into a method of calculating *LD* spectra with the aim of quantitatively analyzing the geometric information yielded by *LD*. This paper relates back to the early work AR undertook on *CD* calculations for DNA in 1992 and also her early *LD* theory work undertaken in 1990.

R2: "Synthetic metallomolecules as agents for the control of DNA structure": This is another review which focuses on how metallo molecules influence DNA structure. It collates much of AR's work in this area with new insights.

R3: "Looking at long molecules in solution: what happens when they are subjected to Couette flow?": This review paper emphasizes the *LD* progress that has been made by AR in the last 5 or so years and also gives new data on actin a fibrous protein.

6: "The binding of single stranded DNA and PNA to single walled carbon nanotubes probed by flow linear dichroism": Carbon nanotubes have exciting potential applications but out ability to charcaterise the properties in solution is fairly limited. This paper reports the first use of *LD* for this purpose with a particular focus on ligand binding. It has lead to a new line of research.

7. "Micro volume Couette flow sample orientation for absorbance and fluorescence linear dichroism": As noted above this reports the invention of the microvolume Couette flow cell which is now enabling the extensive programme on fibrous proteins and membrane proteins.

3. Publications for DSc

REVIEWS: 2003–2007

R1. Bulheller, B.; Rodger, A.; Hirst, J.D. "Circular and linear dichroism of proteins" *Physical Chemistry Chemical Physics*, **2007**, *9*, 2020–2035 — 30%

R2. Richards, A.D.; Rodger, A. "Synthetic metallomolecules as agents for the control of DNA structure" *Chemical Society Reviews,* **2007**, *36*, 471–483 — 50%

R3. Rodger, A.; Marrington, R.; Geeves, M.A.; Hicks, M.; de Alwis, L.; Halsall, D.J.; Dafforn, T.R. "Looking at long molecules in solution: what happens when they are subjected to Couette flow?" *Physical Chemistry Chemical Physics,* **2006**, *8*, 3161−3171 — 80%

R4. Marrington, R.; Seymour, M.; Rodger, A. "A new method for fibrous protein analysis illustrated by application to tubulin microtubule polymerisation and depolymerisation" *Chirality* **2006**, *18*, 680−690 — 70%

R5. Rodger, A.; Marrington, R.; Roper, D.; and Windsor, S. "Circular dichroism spectroscopy for the study of protein ligand interactions" in *Protein–Ligand Interactions: Methods and Protocols,* Uli Nienhaus (ed) in *Methods in Molecular Biology / Methods in Molecular Medicine* The Humana Press Inc., New Jersey, **2005**, 341−361. — 90%

R6. Dafforn, T.R. and Rodger, A. "Unravelling the configuration of protein fibres and membrane proteins" *Current Opinion in Structural Biology*, **2004**, *14*, 541−546 — 50%

R7. Hannon, M.J. and Rodger, A. "Groovy DNA recognition: new agents for gene regulation?" *Pharmaceutical Visions*, **2003**, 14–16 — 30%

INNOVATIVE PAPERS WHICH ESTABLISH NEW LINES OF WORK: 2003–2007

1. Rajendra, J.; Damianoglou, A.; Hicks, M.; Booth, P.; Rodger, P.M.; Rodger, A. "Quantitation of protein orientation in flow-oriented unilamellar liposomes by linear dichroism", *Chemical Physics* **2006**, *326*, 210–220 — 50%

2. Hicks, M.R.; Rodger, A.; Thomas, C.M.; Batt, S.M.; Dafforn, T.R. "Restriction enzyme kinetics monitored by UV linear dichroism" *Biochemistry* **2006**, *45*, 8912–8917 — 20%

3. Marrington,R.; Dafforn, T.R.; Halsall, D.J.; Hicks, M.; Rodger, A. "Validation of new microvolume Couette flow linear dichroism cells" *Analyst*, **2005**, *130*, 1608–1616 — 50%

4. Dafforn, T.R.; Rajendra, J.; Halsall, D.J.; Serpell, L.C.; Rodger, A. "Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell" *Biophysical Journal* **2004**, *86*, 404–410 — 40%

5. Rajendra, J.; Rodger, A. "The binding of single stranded DNA and PNA to single walled carbon nanotubes probed by flow linear dichroism" *Chemistry - A European Journal* **2005**, *11*, 4841–48 — 70%

6. Rajendra, J.; Baxendale, M.; Dit Rap, L.G.; Rodger, A. "Flow linear dichroism to probe binding of aromatic molecules and DNA to single walled carbon nanotubes" *Journal of the American Chemical Society,* **2004**, *126,* 11182−11188 — 60%

7. Marrington, R.; Dafforn, T. R.; Halsall, D.J.; Rodger, A. "Micro volume Couette flow sample orientation for absorbance and fluorescence linear dichroism" *Biophysical Journal*, **2004**, *87*, 2002–2012 — 40%

PAPERS ORIGINATING BECAUSE EXTERNAL SCIENTISTS APPROACHED ME TO COLLABORATE: 2003–2007

8. Shelton, A.; Rodger, A.; McMillin, D. "DNA-Binding Studies of a New Di-Cationic Porphyrin. Insights into Interligand Interactions" *Biochemistry*, **2007**	10%
9. Small, E.; Marrington, R.; Rodger, A.; Scott, D.J.; Sloan K.; Roper, D.; Dafforn, T.R.; Addinall, S.G. "FtsZ polymer-bundling by the *Escherichia coli* ZapA orthologue, YgfE involves a conformational change in bound GTP" *Journal of Molecular Biology*, **2007**	20%
10. Galindo, M.A.; Olea, D.; Romero, M. A.; Gómez, J.; del Castillo, P.; Hannon, M.J.; Rodger, A.; Zamora, F.; Navarro, J.A.R; "Design and Non-covalent DNA Binding of Platinum(II) Metallacalix[n]arenas" *Chemistry — A European Journal*, **2007**	10%
11. Lee, S.C.; Stoilova-Mcphie, S.; Baxter, L.; Fülöp, V.; Henderson, J.; Rodger, A.; Roper, D.; Scott, D.J.; Smith, C.; Morgan, J.A.W; "Biophysical characterisation of an insecticidal toxin component XptA1 suggest a new toxin mode of action" *Journal of Molecular Biology*, **2007**, *366*, 1558–1568	10%
12. Duff, M.R; Tan, W.B.; Bhambhani, A.; Perrin, B.S. Jr.; Thota, J.; Rodger, A.; Kumar, C.V. "Contributions of hydroxyethyl groups to the DNA binding affinities of anthracene probes" *The Journal of Physical Chemistry B* **2006**, *110*, 20693–20701	20%
13. Stootman, F.H.; Fisher, D.M.; Rodger, A.; Aldrich-Wright, J.R. "Improved curve fitting procedures to determine equilibrium binding constants" *The Analyst*, **2006**, *131*, 1145–1151	20%
14. Tan, W. B., Bhambhani, A., Duff, M. R., Rodger, A. and Kumar, C. V. (2006) Spectroscopic identification of binding modes of anthracene probes and DNA sequence recognition. *Photochemistry and Photobiology* **2006**, *82*, 20–30	20%
15. Addinall, S. G.; Johnson, K.A.; Dafforn, T.R.; Smith, C.; Rodger, A.; Gomez, R.P.; Sloan, K.; Blewitt, A.; Scott, D.J.; Roper, D.I "Expression, purification and crystallisation of the cell-division protein YgfE from *Escherichia Coli*" *Acta Cryst. Structural Biology and Crystallisation Communications* **2005,** *F61*, 305–307	5%
16. Uerpmann, C.; Malina, J.; Pascu, M.; Clarkson, G.J.; Moreno, V.; Rodger, A.; Grandas, A.; Hannon, M.J. "Design and DNA binding of an extended triple-stranded metallo-supramolecular cylinder" *Chemistry - A European Journal* **2005**, *11*, 1750–1756	5%
17. Halsall, D.J.; Dafforn, T.R; Marrington, R.; Halligan, E.; and Rodger, A. "Linear Dichroism for quantitative Polymerase Chain Reaction applications: An intrinsic signal from DNA has the potential to be an interference-free, specific and label free probe for real-time and quantitative PCR" *IVD Technology*, **2004**, *6*, 51–60	30%
18. Marrington, R.; Small, E.; Rodger, A.; Dafforn, T.R.; Addinall, S. "FtsZ fibre bundling is triggered by a calcium-induced conformational change in bound GTP" *Journal of Biological Chemistry* **2004**, *47*, 48821–48829	20%
19. Ambroż, H.B.; Kemp, T.J.; Rodger, A.; and Przybytniak, G. "Ferric and ferrous ions: binding to DNA and influence on radiation-induced processes" *Radiation Physics and Chemistry*, **2004**, *71*, 1023–1030	10%

20. McCabe, R.W., Rodger, A. and Taylor, A. "A study of the secondary structure of *Candida antarctica* lipase B using synchrotron radiation circular dichroism measurements" *Enzyme and Microbial Technology*, **2004**, *36*, 70–74 — 10%

21. Glover, P.B.; Ashton, P.R.; Childs, L.J.; Rodger, A.; Kercher, M., Williams, R.M., Luisa De Cola, L. and Pikramenou, Z., "Hairpin-Shaped Heterometallic Luminescent Lanthanide Complexes for DNA Intercalative Recognition" *Journal of the American Chemical Society* **2003**, *125*, 9918–9919 — 5%

22. Miles, A.J., Wien, F., Lees, J.G., Rodger, A., Janes, R.W. and Wallace, B.A. "Calibration and Standardisation of Synchrotron Radiation Circular Dichroism and Conventional Circular Dichroism Spectrophotometers" *Spectroscopy* **2003**, 17, 653–661 — 10%

23. Novakova, O., Chen, H., Vrana, O., Rodger, A., Sadler, P.J., Brabec, V. "DNA interactions of monofunctional organometallic ruthenium(II) antitumor complexes in cell-free media" *Biochemistry* **2003,** *42*, 11544–11554 — 5%

24. Đapic, V., Abdomerovic, V., Marrington, R., Peberdy, J.C., Rodger, A., Trent, J.O., Bates, P.J., "Biophysical and biological properties of G-quartet forming oligonucleotides" *Nucleic Acids Research* **2003**, *31*, 2097–2107 — 10%

25. Grummitt, A.R., Harding, M.M., Anderberg, P.I., Rodger, A. "Carbohydrate derivatives of the antitumour alkaloid 9-hydroxyellipticine" *European Journal of Organic Chemistry* **2003**, *1*, 63–71 — 20%

WARWICK COLLABORATIONS: 2003–2007

26. Peberdy, J.C.; Malina, J.; Khalid, S.; Hannon, M.J.; Rodger, A. "Influence of surface shape on DNA binding of bimetallo helicates" *Journal of Inorganic Biochemistry* **2007,** — 60%

27. Kerckhoffs, J.M.C.A.; Peberdy, J.C.; Meistermann, I.; Childs, L.J.; Isaac, C.J.; Pearmund, C.R.; Reudegger, V.; Alcock, N.W.; Hannon, M.J.; Rodger, A. "Enantiomeric resolution of supramolecular helicates with different surface topographies" *Dalton Transactions*, **2007**, 734–742 — 30%

28. Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. "Shape effects on the activity of synthetic major-groove binding ligands" *Journal of Molecular Graphics and Modelling*, **2007**, *25*, 794–800 — 10%

29. Hannon, M.J.; Green, P.S.; Fisher, D.M.; Derrick, P.J.; Beck, J.L.; Watt, S.J.; Ralph, S.F.; Shiel, M.M.; Barker, P.R.; Alcock, N.W.; Price, R.J.; Sanders, K.J.; Pither, R.; Davis, J.; Rodger, A. "An estrogen-platinum terpyridine conjugate: DNA and protein binding and cellular delivery", *Chemistry — A European Journal*, **2006,** *12*, 8000-8013 — 60%

30. Childs, L.J.; Malina, J.; Pascu, M.; Rolfsnes, B.E.; Prieto, M.J.; Broome, M.J.; Rodger, P.M.; Sletten, E.; Moreno, V.; Rodger, A.; Hannon, M.J. "A DNA-binding cuprous metallo-supramolecular cylinder that can act as an artificial nuclease" *Chemistry — A European Journal* **2006**, *12*, 4919–4927 — 10%

31. Oleksi, A.; Blanco, A.G.; Boer, R.; Usón, I.; Aymamí, J.; Rodger, A.; Hannon, M.J.; Coll, M. "Molecular recognition of a three-way DNA junction by a metallo-supramolecular helicate", *Angewandte Chemie* **2006**, 45,1227– — 10%

32. Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. "Simulations of DNA coiling around a synthetic supramolecular cylinder that binds in the — 10%

DNA major groove" *Chemistry —A European Journal* **2006**, *12*, 3493-506

33. Khalid, S.; Rodger, P.M.; and Rodger,A. "Theoretical aspects of the enantiomeric resolution of dimetallo helicates with different surface topologies on cellulose columns" *Journal of Liquid Chromatography* **2005**, *28*, 2995–3003

34. Koutsoudaki, C.; Kresk, M.; Rodger, A. "Chemical composition and antibacterial activity of the essential oil and the gum of *Pistacia lentiscus var. chia*" *Journal of Agricultural and Food Chemistry* **2005**, *53*, 7681−7685 30%

35. Alcock, N.W., Sanders, K.J., Rodger, A. "Potential injectable contraceptive steroids: testosterone buciclate" *Acta Crystallographica* **2004**, *E60*, o348–o349 50%

36. Miguel, M.S., Marrington, R., Rodger, P.M., Rodger, A., and Robinson, C. "An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on hydrophobicity of the environment" *European Journal of Biochemistry* **2003**, *270*, 3345–3352 15%

Appendix 1

Alison Rodger: Complete publication record in reverse chronological order

Theses

T4. Rodger, A. "Molecular aspects of biomolecule structure and function". *DSc Thesis, Sydney,* **2002**

T3. Rodger, A. "Doctoral thesis abstract: Symmetry selection rules: Analytic development and chemical application" *Journal and Proceedings, Royal Society of New South Wales* **1986**, *119*, 141

T2. Rodger, A. "Symmetry selection rules: Analytic development and chemical application", *PhD Thesis, Sydney,* **1985**

T1. Rodger, A. "Coupled oscillator circular dichroism of inclusion complexes", *Honours Thesis, Sydney,* **1981**

Books

B6. Lawrence, C. H.; Rodger, A.; Compton, R. G. "Foundations of physical chemistry"; **2005**

B5. Lawrence, C.H, Rodger, A, Compton, R.G. "Fundamentos de química física" Traductor: Calvo, E.J.; *Eudeba, Buenos Aires, by arrangement with Oxford University Press, Oxford,* **2000**

B4. Lawrence, C. H.; Rodger, A.; Compton, R. G. "Foundations of physical chemistry"; *Oxford University Press: Oxford, Japanese Edition* **1998**

B3. Rodger, A.; Nordén, B. "Circular dichroism and linear dichroism"; *Oxford University Press*, **1997**, pp150

B2. Lawrence, C. H.; Rodger, A.; Compton, R. G. "Foundations of physical chemistry"; *Oxford University Press: Oxford,* **1996,** pp 96

B1. Rodger, A.; Rodger, P. M. "Molecular geometry"; *Butterworth-Heineman Ltd.: Oxford,* **1995**, pp 190

Refereed book chapters

C9. Rodger, A.; Marrington, R.; Roper, D.; and Windsor, S. "Circular dichroism spectroscopy for the study of protein ligand interactions" in *Protein–Ligand Interactions: Methods and Protocols,* Uli Nienhaus (ed) in *Methods in Molecular Biology / Methods in Molecular Medicine* The Humana Press Inc., New Jersey, **2005**, 341−361.

C8. Rodger, A., Rajendra, J., Marrington, R., Mortimer, R., Andrews, T., Hirst, J.B., Gilbert, A.T.B., Halsall, D., Dafforn, T., Ardhammar, M., Nordén, B., Woolhead, C.A., Robinson, C., Pinheiro, T., Kazlauskaite J., Seymour, M., Perez, N., Hannon, M.J. "Flow oriented linear dichroism to probe protein orientation in membrane environments" in *Biophysical Chemistry: Membranes and Proteins,* Templer, R.H. and Leatherbarrow, R. (eds) The Royal Society of Chemistry, Cambridge, **2002**, pp 3−19

C7. Rodger, A. "Circular dichroism and linear dichroism" in *Encyclopedia of analytical chemistry: instrumentation and applications* Meyers R.A. (ed) John Wiley and Sons **2000**, 30 pages

C6. Rodger, A., Carey, M. "Stopped-flow circular dichroism" in *Spectrometry and*

spectrofluorimetry: A practical approach Gore, M. (Ed.) **2000**, pp 265−281

C5. Rodger, A., Ismail, M.A. "Introduction to circular dichroism" in *Spectrometry and spectrofluorimetry: A practical approach* Gore, M. (Ed.) **2000**, pp 99−139

C4. Rodger, A., Sanders, K.J. "Biomacromolecular applications of UV-visible absorption" in *Encyclopedia of spectroscopy and spectrometry, Academic Press*, **1999**, 130−139

C3. Rodger, A. "Linear dichroism" In *Methods in Enzymology*; J. F. Riordan and B. L. Vallee, Eds.; Academic Press: San Diego, **1993**; Vol. 226; pp 232−258

C2. Johnson, B. F. G.; Bott, A.; Benfield, R. E.; Braga, D.; Marseglia, E. A.; Rodger, A. "Mechanistic features of carbonyl cluster rearrangement" In *IUCCP Symposium: Metal-Metal Bonds and Clusters in Chemistry and Catalysis*; Plenum: New York, **1990**; pp 141−160

C1. Johnson, B.F.G.; Rodger, A. In "Polyhedral rearrangements and fragmentation reactions in clusters"; D. F. Schriver, H. Kaesz and R. Adams, Eds.; VCH Verlagsgellschaft mgH: Federal Republic of Germany, **1990**; pp 303−327

Papers

124. Kant, S.; Mitchell, G; Walsh, J.; Hall, C.R.; Rodger, A. "An introduction to communicating science" *Higher Education Academy 'Physical Sciences' online journal New Directions* **2007**
123. Shelton, A.; Rodger, A.; McMillin, D. "DNA-Binding Studies of a New Di-Cationic Porphyrin. Insights into Interligand Interactions" *Biochemistry*, **2007**
122. Peberdy, J.C.; Malina, J.; Khalid, S.; Hannon, M.J.; Rodger, A. "Influence of surface shape on DNA binding of bimetallo helicates" *Journal of Inorganic Biochemistry* **2007,**
121. Bulheller, B.; Rodger, A.; Hirst, J.D. "Circular and linear dichroism of proteins" *Physical Chemistry Chemical Physics***, 2007**, *9*, 2020−2035
120. Small, E.; Marrington, R.; Rodger, A.; Scott, D.J.; Sloan· K.; Roper, D.; Dafforn, T.R.; Addinall, S.G. "FtsZ polymer-bundling by the *Escherichia coli* ZapA orthologue, YgfE involves a conformational change in bound GTP" *Journal of Molecular Biology*, **2007**
119. Galindo, M.A.; Olea, D.; Romero, M. A.; Gómez, J.; del Castillo, P.; Hannon, M.J.; Rodger, A.; Zamora, F.; Navarro, J.A.R; "Design and Non-covalent DNA Binding of Platinum(II) Metallacalix[n]arenas" *Chemistry —A European Journal*, **2007**
118. Kerckhoffs, J.M.C.A.; Peberdy, J.C.; Meistermann, I.; Childs, L.J.; Isaac, C.J.; Pearmund, C.R.; Reudegger, V.; Alcock, N.W.; Hannon, M.J.; Rodger, A. "Enantiomeric resolution of supramolecular helicates with different surface topographies" *Dalton Transactions*, **2007**, 734−742
117. Lee, S.C.; Stoilova-Mcphie, S.; Baxter, L.; Fülöp, V.; Henderson, J.; Rodger, A.; Roper, D.; Scott, D.J.; Smith, C.; Morgan, J.A.W; "Biophysical characterisation of an insecticidal toxin component XptA1 suggest a new toxin mode of action" *Journal of Molecular Biology*, **2007**, *366*, 1558−1568
116. Richards, A.D.; Rodger, A. "Synthetic metallomolecules as agents for the control of DNA structure" *Chemical Society Reviews*, **2007**, *36*, 471−483
115. Duff, M.R; Tan, W.B.; Bhambhani, A.; Perrin, B.S. Jr.; Thota, J.; Rodger, A.; Kumar, C.V. "Contributions of hydroxyethyl groups to the DNA binding affinities of anthracene probes" *The Journal of Physical Chemistry B* **2006**, *110*, 20693–20701
114. Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. "Shape effects on the activity of synthetic major-groove binding ligands" *Journal of Molecular Graphics and Modelling*, **2007**, *25*, 794−800
113. Stootman, F.H.; Fisher, D.M.; Rodger, A.; Aldrich-Wright, J.R. "Improved curve fitting procedures to determine equilibrium binding constants" *The Analyst*, **2006**, *131*, 1145−1151
112. Rodger, A.; Marrington, R.; Geeves, M.A.; Hicks, M.; de Alwis, L.; Halsall, D.J.;

Dafforn, T.R. “Looking at long molecules in solution: what happens when they are subjected to Couette flow?” *Physical Chemistry Chemical Physics*, **2006**, *8*, 3161−3171

111. Hannon, M.J.; Green, P.S.; Fisher, D.M.; Derrick, P.J.; Beck, J.L.; Watt, S.J.; Ralph, S.F.; Shiel, M.M.; Barker, P.R.; Alcock, N.W.; Price, R.J.; Sanders, K.J.; Pither, R.; Davis, J.; Rodger, A. “An estrogen-platinum terpyridine conjugate: DNA and protein binding and cellular delivery”, *Chemistry —A European Journal*, **2006,** *12*, 8000-8013
110. Hicks, M.R.; Rodger, A.; Thomas, C.M.; Batt, S.M.; Dafforn, T.R. “Restriction enzyme kinetics monitored by UV linear dichroism” *Biochemistry* **2006**, *45*, 8912−8917
109. Marrington, R.; Seymour, M.; Rodger, A. “A new method for fibrous protein analysis illustrated by application to tubulin microtubule polymerisation and depolymerisation” *Chirality* **2006**, *18*, 680–690
108. Childs, L.J.; Malina, J.; Pascu, M.; Rolfsnes, B.E.; Prieto, M.J.; Broome, M.J.; Rodger, P.M.; Sletten, E.; Moreno, V.; Rodger, A.; Hannon, M.J. “A DNA-binding cuprous metallo-supramolecular cylinder that can act as an artificial nuclease” *Chemistry —A European Journal* **2006**, *12*, 4919–4927
107. Oleksi, A.; Blanco, A.G.; Boer, R.; Usón, I.; Aymamí, J.; Rodger, A.; Hannon, M.J.; Coll, M. “Molecular recognition of a three-way DNA junction by a metallo-supramolecular helicate”, *Angewandte Chemie* **2006**, 45,1227–
106. Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. “Simulations of DNA coiling around a synthetic supramolecular cylinder that binds in the DNA major groove” *Chemistry —A European Journal* **2006**, *12*, 3493-506
105. Rajendra, J.; Damianoglou, A.; Hicks, M.; Booth, P.; Rodger, P.M.; Rodger, A. “Quantitation of protein orientation in flow-oriented unilamellar liposomes by linear dichroism”, *Chemical Physics* **2006**, *326*, 210–220
104. Tan, W. B., Bhambhani, A., Duff, M. R., Rodger, A. and Kumar, C. V. (2006) Spectroscopic identification of binding modes of anthracene probes and DNA sequence recognition. *Photochemistry and Photobiology* **2006**, *82*, 20–30
103. Marrington,R.; Dafforn, T.R.; Halsall, D.J.; Hicks, M.; Rodger, A. “Validation of new microvolume Couette flow linear dichroism cells” *Analyst*, **2005**, *130*, 1608−1616
102. Khalid, S.; Rodger, P.M.; and Rodger,A. “Theoretical aspects of the enantiomeric resolution of dimetallo helicates with different surface topologies on cellulose columns” *Journal of Liquid Chromatography* **2005**, *28*, 2995–3003
101. Koutsoudaki, C.; Kresk, M.; Rodger, A. “Chemical composition and antibacterial activity of the essential oil and the gum of *Pistacia lentiscus var. chia*” *Journal of Agricultural and Food Chemistry* **2005**, *53*, 7681−7685
100. Rajendra, J.; Rodger, A. “The binding of single stranded DNA and PNA to single walled carbon nanotubes probed by flow linear dichroism” *Chemistry - A European Journal* **2005**, *11*, 4841−48
99. Addinall, S. G.; Johnson, K.A.; Dafforn, T.R.; Smith, C.; Rodger, A.; Gomez, R.P.; Sloan, K.; Blewitt, A.; Scott, D.J.; Roper, D.I “Expression, purification and crystallisation of the cell-division protein YgfE from *Escherichia Coli*” *Acta Cryst. Structural Biology and Crystallisation Communications* **2005,** *F61*, 305−307
98. Uerpmann, C.; Malina, J.; Pascu, M.; Clarkson, G.J.; Moreno, V.; Rodger, A.; Grandas, A.; Hannon, M.J. “Design and DNA binding of an extended triple-stranded metallo-supramolecular cylinder” *Chemistry - A European Journal* **2005**, *11*, 1750–1756
97. Halsall, D.J.; Dafforn, T.R; Marrington, R.; Halligan, E.; and Rodger, A. “Linear Dichroism for quantitative Polymerase Chain Reaction applications: An intrinsic signal from DNA has the potential to be an interference-free, specific and label free probe for real-time and quantitative PCR” *IVD Technology*, **2004**, *6*, 51−60
96. Dafforn, T.R. and Rodger, A. “Unravelling the configuration of protein fibres and membrane proteins” *Current Opinion in Structural Biology*, **2004**, *14*, 541−546
95. Marrington, R.; Small, E.; Rodger, A.; Dafforn, T.R.; Addinall, S. “FtsZ fibre bundling

is triggered by a calcium-induced conformational change in bound GTP" *Journal of Biological Chemistry* **2004**, *47*, 48821−48829

94. Rajendra, J.; Baxendale, M.; Dit Rap, L.G.; Rodger, A. "Flow linear dichroism to probe binding of aromatic molecules and DNA to single walled carbon nanotubes" *Journal of the American Chemical Society*, **2004**, *126*, 11182−11188
93. Ambroż, H.B.; Kemp, T.J.; Rodger, A.; and Przybytniak, G. "Ferric and ferrous ions: binding to DNA and influence on radiation-induced processes" *Radiation Physics and Chemistry*, **2004**, *71*, 1023−1030
92. Marrington, R.; Dafforn, T. R.; Halsall, D.J.; Rodger, A. "Micro volume couette flow sample orientation for absorbance and fluorescence linear dichroism" *Biophysical Journal*, **2004**, *87*, 2002−2012
91. McCabe, R.W., Rodger, A. and Taylor, A. "A study of the secondary structure of *Candida antarctica* lipase B using synchrotron radiation circular dichroism measurements" *Enzyme and Microbial Technology*, **2004**, *36*, 70−74
90. Alcock, N.W., Sanders, K.J., Rodger, A. "Potential injectable contraceptive steroids: testosterone buciclate" *Acta Crystallographica* **2004**, *E60*, o348−o349
89. Dafforn, T.R.; Rajendra, J.; Halsall, D.J.; Serpell, L.C.; Rodger, A. "Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell" *Biophysical Journal* **2004**, *86*, 404−410
88. Glover, P.B.; Ashton, P.R.; Childs, L.J.; Rodger, A.; Kercher, M., Williams, R.M., Luisa De Cola, L. and Pikramenou, Z., "Hairpin-Shaped Heterometallic Luminescent Lanthanide Complexes for DNA Intercalative Recognition" *Journal of the American Chemical Society* **2003**, *125*, 9918–9919
87. Hannon, M.J. and Rodger, A. "Groovy DNA recognition: new agents for gene regulation?" *Pharmaceutical Visions*, **2003**, 14−16
86. Miguel, M.S., Marrington, R., Rodger, P.M., Rodger, A., and Robinson, C. "An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on hydrophobicity of the environment" *European Journal of Biochemistry* **2003**, *270*, 3345−3352
85. Miles, A.J., Wien, F., Lees, J.G., Rodger, A., Janes, R.W. and Wallace, B.A. "Calibration and Standardisation of Synchrotron Radiation Circular Dichroism and Conventional Circular Dichroism Spectrophotometers" *Spectroscopy* **2003**, 17, 653−661
84. Novakova, O., Chen, H., Vrana, O., Rodger, A., Sadler, P.J., Brabec, V. "DNA interactions of monofunctional organometallic ruthenium(II) antitumor complexes in cell-free media" *Biochemistry* **2003,** *42*, 11544−11554
83. Đapic, V., Abdomerovic, V., Marrington, R., Peberdy, J.C., Rodger, A., Trent, J.O., Bates, P.J., "Biophysical and biological properties of G-quartet forming oligonucleotides" *Nucleic Acids Research* **2003**, *31*, 2097–2107
82. Grummitt, A.R., Harding, M.M., Anderberg, P.I., Rodger, A. "Carbohydrate derivatives of the antitumour alkaloid 9-hydroxyellipticine" *European Journal of Organic Chemistry* **2003**, *1*, 63−71
81. Rodger, A., Patel, K.K., Sanders, K.J., Datt, M., Sacht, C., Hannon, M.J. "Anti-tumour platinum acylthiourea complexes and their interactions with DNA" *Dalton* **2002**, 3656–3663
80. Rodger, A., Rajendra, J., Marrington, R., Ardhammar, M.,Nordén, B., Hirst, J.D., Gilbert, A.T.B. Dafforn, T.R., Halsall, D.J., Woolhead, C.A., Robinson, C. Pinheiro, T.J., Kazlauskaite, J., Seymour, M., Perez, N., Hannon, M.J. "Flow oriented linear dichroism to probe protein orientation in membrane environments" *Phys. Chem. Chem. Phys.* **2002**, *4*, 4051–4057
79. Dharmagunawardena, B., Sanders, K.J., Cannan, S., Howell, K., Ilchyshyn, A., Rodger, A. "Gas chromatography/mass spectrometry as an investigative tool in multiple allergy to essential oils" *Contact Dermatitis* **2002,** *47*, 288–292

78. Lukač, R., Clark, A.J., San-Miguel, M.A., Rodger, A., Rodger, P.M. "Predicting chiral discrimination in HPLC from computer simulations" *Journal of Molecular Liquids* **2002**, *101*, 261–272
77. Pinheiro, P., Scarlett, G., Rodger, A., Rodger, P.M., Murray, A., Brown, T., Newbury, S., McClellan, J. "Structures of CUG repeats: potential implications for human genetic diseases" *Journal of Biological Chemistry* **2002**, 277, 35183–35190
76. Đapic, V., Bates, P.J., Trent, J.O., Rodger, A., Thomas, S.D., Miller, D.M., "Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications" *Biochemistry* **2002**, *41*, 3676–3685
75. Moldeheim, E., Hannon, M.J., Meistermann, I., Rodger, A., Sletten, E., "Interaction between a DNA oligonucleotide and dinuclear iron(II) supramolecular cylinder: an NMR and molecular dynamics study" *Journal of Biological Inorganic Chemistry* **2002**, *7*, 770–780
74. Ardhammar, M., Lincoln, P., Rodger, A., Nordén, B. "Absolute configuration and assignment of electronic transitions of light-switch complex $[Ru(phen)_2dppz]^{2+}$ deduced from oriented circular dichroism in a lamellar liquid crystal host" *Chemical Physics Letters* **2002**,*354*, 44–50
73. Meistermann, I., Moreno, V., Prieto, M.J., Molderheim, E., Sletten, E., Khalid, S., Rodger, P.M., Peberdy, J., Isaac, C.J., Rodger, A., Hannon, M.J. "Intramolecular DNA coiling mediated by metallo-supramolecular cylinders: differential binding of P and M helical enantiomers" *Proceedings of the National Academy of Sciences*, **2002**, 99, 5069–5074
72. Patel, K.K., Plummer E.A., Darwish, M., Rodger, A., Hannon, M.J. "Aryl substituted ruthenium bisterpyridine complexes: intercalation and groove binding with DNA" *J. Inorganic Biochemistry* **2002**, *91*, 220–229
71. Lukač, R., Clark, A.J., Khalid, S., Rodger, A., Snedden, A., Rodger, P.M. "Chiral discrimination in mobile phases for HPLC" *Journal of Molecular Liquids* **2002**, *98–99*, 411–423
70. Rodger, A.; Nordén, B.; Rodger, P.M.; Bates, P.J. "DNA as a catalyst and catalytic template for the racemisation of metal tris-phenanthroline complexes" *European Journal of Inorganic Chemistry* **2002**, *1*, 49–53
69. Dafforn, T.R.; Halsall, D.J.; Rodger, A. "The detection of single base pair mutations using a novel spectroscopic technique" *Chemical Communications* **2001**, 2410–2411
68. Woolhead, C.A.; Thompson, S.J.; Moore, M.; Tissier, C.; Mant, A.; Rodger, A.; Henry, R.; Robinson, C. "Distinct albino3-dependent and -independent pathways for thylakoid membrane protein insertion" *Journal of Biological Chemistry* **2001**, *276*, 40841–40846
67. Hannon, M.J., Meistermann, I., Isaac, C.J., Blomme, C., Aldrich-Wilson, J.R., Rodger, A. "Paper: a cheap yet effective chiral stationary phase for chromatographic resolution of metallo-supramolecular helicates" *Chemical Communications* **2001**, 1078–1079
66. Rodger, A.; Parkinson, A. and Best, S. "Molecular features of Co(III) tetra and pentammines affect their influence on DNA structure" *European Journal of Inorganic Chemistry*, **2001**, *9*, 2311–2316
65. Jackson, A., Davis, J., Pither, R.J., Rodger, A., and Hannon, M.J. "Estrogen-derived steroidal metal complexes: Agents for cellular delivery of metal centers to estrogen receptor-positive cells" *Inorganic Chemistry*, **2001**, *40*, 3964–3973
64. Etherington, K.J. Rodger, A.; Hemming, P. "CHN microanalysis a technique for the 21st century?" *LabPlus International*, **2001**, *Feb/March*, 26–27
63. Scigelova, M., Green, P.S., Giannakopulos, A.E., Rodger, A., Crout, D.H.G.C., Derrick, P.J. "A practical protocol for reduction of disulfide bonds in proteins prior to analysis by mass spectrometry" *European Journal of Mass Spectrometry*, **2001**, *7*, 29–34
62. Woolhead, C.A., Mant, A., Kim, S.J., Robinson, C., Rodger, A. "Conformation of a purified, 'spontaneously' inserting thylakoid membrane protein precursor in aqueous

solvent and detergent micelles" *Journal of Biological Chemistry*, **2001**, *276*, 14607–14613

61. Sanders, K.J., Wiles, H., Rodger, A. "Dialysis cells for controlled DNA:drug binding studies" *The Analyst*, **2001**, *126*, 852–854
60. Hannon, M.J., Moreno, V., Prieto, M.J., Molderheim, E., Sletten, E., Meistermann, I., Isaac, C.J., Sanders, K.J., Rodger, A. "Intramolecular DNA coiling mediated by a metallo supramolecular cylinder" *Angewandte Chemie*, **2001**, *40*, 879–884
59. Parkinson, A., Hawken, M., Hall, M., Sanders, K.J., Rodger, A. "Amine induced Z-DNA in poly(dG-dC).(dG-dC): circular dichroism and gel electrophoresis study" *PCCP*, **2000**, *2*, 5469–5478
58. Pandya, M.J., Spooner, G.M., Sunde, M., Thorpe, J.R., Rodger, A., Woolfson, D.N. "Sticky-end assembly of a designed protein fibre provides insight into protein fibrillogenesis" *Biochemistry*, **2000**, *39*, 8728–8734
57. Rodger, A., Sanders, K.J., Hannon, M.J., Meistermann, I., Parkinson, A., Vidler, D.S., Haworth, I.S. "DNA structure control by polycationic species: polyamines, cobalt ammines, and di-metallo transition metal chelates" *Chirality*, **2000**, *12*, 221–236
56. Ismail, M.A., Rodger, P.M., Rodger, A. "Drug self-assembly on DNA: sequence effects with *trans*-*bis*-*N*-methylpyridiniumyl)diphenyl porphyrin and Hoechst 33258" *Journal of Biomolecular Structure and Dynamics*, **2000**, *S2*, 335–348
55. Coggan, D. Z., Haworth, I. S., Bates, P. J., Robinson, A., Rodger, A.; "DNA binding of ruthenium *tris*-(1,10-phenanthroline): evidence for the dependence of binding mode on metal complex concentration" *Inorganic Chemistry*, **1999**, *38*, 4486–4497
54. Rodger, A., Nordén, B. "A coupled-oscillator type approach to magnetic circular dichroism" *Enantiomer*, **1998**, *3*, 409–421
53. Rodger, A.; Latham, H.C.; Wormell, P.; Parkinson, A.; Ismail, M.; Sanders, K.J. "DNA-drug system: how circular dichroism data is complemented by other spectroscopic techniques" *Enantiomer*, **1998**, *3*, 395–408
52. Ismail, M.A.; Sanders, K.J.; Fennel, G.C.; Latham, H.C.; Wormell, P.; Rodger, A. "Spectroscopic studies of 9-Hydroxyellipticine binding to DNA" *Biopolymers*, **1998**, *46*, 127–143
51. Harding, M.M.; Krippner, G.Y.; Shelton, C.J.; Rodger, A.; Sanders, K.J.; Mackay, J.P.; Prakash, A.S. "DNA-binding studies of XSPTSPSZ derivatives of the intercalating heptad repeat of RNA polymerase II" *Biopolymers*, **1997**, *42*, 387–398
50. Kim, S. K.; Sun, J.-S.; Garestier, T.; Hélène, C.; Nguyen, C. H.; Bisagni, E.; Rodger, A.; Nordén, B. "Binding geometries and protonation states of triple helix selective benzopyrido[4,3-b]indole ligands complexed with double and triple helical polynucleotides" *Biopolymers*, **1997**, *42*, 101–111
49. Elcock, A.H.; Rodger, A.; Richards, W.G. "Theoretical studies of the intercalation of 9-hydroxyellipticine in DNA" *Biopolymers* **1996**, *39*, 309–326
48. Newbury, S. F.; Mc Clellan, J. A.; Rodger, A. "Spectroscopic and thermodynamic studies of conformational changes in long, natural mRNA molecules" *Analytical Communications* **1996**, *33*, 117–122
47. Moon, J.-H.; Kim, S. K.; Sehlstedt, U.; Rodger, A.; Nordén, B. "DNA structural features responsible for sequence dependent binding geometries of Hoechst 33258" *Biopolymers* , **1996**, *38*, 593–606
46. Kim, H. K.; Kim, J.-M.; Kim, S. K.; Rodger, A.; Nordén, B. "Interactions of intercalative and minor groove binding ligands with triplex poly(dA).2poly(dT) and with duplex poly(dA).poly(dT) and poly[d(A-T)$_2$] studied by CD, LD and normal absorption" *Biochemistry*, **1996**, *35*, 1187–1194
45. Mc Coubrey, A.; Latham, H. C.; Rodger, A.; Lowe, G. "4-Picoline-2,2':6',2''-terpyridine-platinum(II) - a potent intercalator of DNA" *FEBS Letters*, **1996**, *380*, 73–78

44. Rodger, A.; Taylor, S.; Adlam, G.; Blagbrough, I. S.; Haworth., I. S. "Multiple DNA binding modes of anthracene-9-carbonyl-N^1-spermine" *Bioorganic and Medicinal Chemistry* **1995**, *3*, 861−872.
43. Baldwin, J. E.; Claridge, T. D.; Hulme, C.; Rodger, A.; Schofield, C. J. "Comments on the use of a dichromophoric circular dichroism assay for the identification of beta-turns in peptides" *International Journal of Peptide and Protein Research* **1994**, *43*, 180−183
42. Fidler, J.; Rodger, P. M.; Rodger, A. "Chiral solvent structure around chiral molecules: experimental and theoretical study" *Journal of the American Chemical Society* **1994**, *116*, 7266−7273
41. Adlam, G.; Blagbrough, I. S.; Taylor, S.; Latham, H. C.; Haworth, I. S.; Rodger, A. "Multiple binding modes with DNA of anthracene-9-carbonyl-N^1-spermine probed by linear dichroism, circular dichroism, normal absorption, and molecular modelling compared with those of spermidine and spermine" *Bioorganic and Medicinal Chemistry Letters* **1994**, *4*, 2435−2440
40. Rodger, A.; Blagbrough, I. S.; Adlam, G.; Carpenter, M. L. "DNA binding of a spermine derivative: spectroscopic study of anthracene-9-carbonyl-N^1-spermine with poly(dG-dC)$_2$ and poly(dA-dT)$_2$" *Biopolymers* **1994**, *34*, 1583–1593
39. Johnson, B. F. G.; Rodger, A.; Colwell, S. M. "A unified view of inorganic stereochemistry and stereochemical changes" *Inorganica Chimica Acta* **1994**, *219*, 187−192
38. Fidler, J.; Rodger, P. M.; Rodger, A. "Circular dichroism as a probe of chiral solvent structure around chiral molecules" *Journal of the Chemical Society Perkin II* **1993**, 235−241
37. Lyng, R.; Rodger, A.; Nordén, B. "The circular dichroism of drug-DNA systems. 2. Poly(dA-dT) B-DNA" *Biopolymers* **1992**, *32*, 1201−1214
36. Haworth, I. S.; Rodger, A.; Richards, W. G. "A molecular dynamics simulation of a polyamine-induced conformational change of DNA: a possible mechanism for the B to Z transition" *Journal of Biomolecular Structure and Dynamics* **1992**, *10*, 195–211
35. Kemp-Harper, R.; Rodger, A.; Compton, R. G. "An experimental insight into electrochemistry" *Education in Chemistry* **1992**, *29*, 163−165
34. Rodger, A.; Johnson, B. F. G. "The significance of ligand-ligand interactions for transition metal complex geometries" *Inorganica Chimica Acta* **1992**, *191*, 109−113
33. Lyng, R.; Rodger, A.; Nordén, B. "The circular dichroism of drug-DNA systems. 1. Poly(dG-dC) B-DNA" *Biopolymers* **1992**, *31*, 1709−1719
32. Haworth, I. S.; Elcock, A.H; Rodger, A.; Richards, W. G. "A binding mode of Λ-[tris(1,10-phenanthroline) ruthenium(II)]$^{2+}$ exhibiting preference for purine-3',5'-pyrimidine sites of DNA" *Journal of Biomolecular Structure and Dynamics* **1991**, *9*, 553−569
31. Rodger, A. "Template symmetry restrictions on reaction mechanisms with application to transition metal complexes" *Inorganica Chimica Acta* **1991**, *195*, 193−200
30. Haworth, I. S.; Elcock, A. H.; Freeman, J.; Rodger, A.; Richards, W. G. "Sequence selective binding to the DNA major groove: Tris(1,10-phenanthroline) metal complexes binding to poly(dG-dC) and poly(dA-dT)" *Journal of Biomolecular Structure and Dynamics* **1991**, *9*, 23−44
29. Rodger, A.; Moloney, M. "n-π* Circular dichroism of planar zig-zag carbonyl compounds" *Journal of the Chemical Society Perkin II* **1991**, 919–925
28. Haworth, I. S.; Rodger, A.; Richards, W. G. "A molecular mechanics study of spermine complexation to DNA: A new model for spermine / poly(dG-dC) binding" *Proceedings of the Royal Society Series B* **1991**, *244*, 107−116
27. Basil, A.; Ben-Tzur, S.; Gedanken, A.; Rodger, A. "An extension of the quadrant rule in oxiranes to non-alkyl substituents: the CD of R(-) and S(+) epichlorohydrin" *Chemical Physics Letters* **1991**, *180*, 482−484

26. Rodger, A.; Gedanken, A.; Klein, H. “Circular dichroism of molecules requiring two substituents for chirality” *Molecular Physics* **1991**, *72*, 803–815
25. Johnson, B. F. G.; Bott, A. W.; Hugh-Jones, D.; Rodger, A. “Geometrical and orientational isomers in clusters” *Polyhedron* **1990**, *9*, 1769–1774
24. Braga, D.; Rodger, A.; Johnson, B. F. G. “A ligand-ligand interaction model for the structures of transition metal clusters” *Inorganica Chimica Acta* **1990**, *174*, 185–191
23. Hiort, C.; Nordén, B.; Rodger, A. “Enantioselective DNA binding of [Ru(1,10-phenanthroline)$_3$]$^{2+}$ studied with linear dichroism” *Journal of the American Chemical Society* **1990**, *112*, 1971–1982
22. Rodger, A.; Colwell, S. M.; Johnson, B. F. G. “Bonding in boranes” *Polyhedron* **1990**, *9*, 1035−1044
21. Rodger, A.; Johnson., B. F. G. “Structure and reactivity of transition metal complexes: A stereochemical approach” *Polyhedron* **1989**, *8*, 1742−1744
20. Johnson, B. F. G.; Rodger, A. “Energetics of isomerisation reactions in transition metal complexes” *Inorganic Chemistry* **1989**, *28*, 1003−1006
19. Rodger, A.; Johnson, B. F. G. “Which is more likely: The Ray Dutt twist or the Bailar twist” *Inorganic Chemistry* **1988**, *27*, 3061−3062
18. Rodger, A.; Johnson, B. F. G. “Geometry determination for AL$_n$ systems: Principles and applications” *Inorganica Chimica Acta* **1988**, *146*, 37–53
17. Rodger, A. “Oxirane circular dichroism: An independent systems / perturbation approach” *Journal of the American Chemical Society* **1988**, *110*, 5941−5945
16. Rodger, A.; Rodger, P. M. “The circular dichroism of the carbonyl n-π* transition: An independent systems / perturbation approach” *Journal of the American Chemical Society* **1988**, *110*, 2361–2368
15. Schipper, P. E.; Rodger, A. “A dispersion induced circular dichroism and normal absorption study of the n-pi* carbonyl transition for the isoelectronic substituent series -CH$_3$, -NH$_2$, -OH” *Spectrochimica Acta* **1988**, *44A*, 575–580
14. Rodger, A.; Johnson, B. F. G. “Polyhedral rearrangements of metal clusters” *Polyhedron* **1988**, *7*, 1107−1120
13. Rodger, A.; Schipper, P. E. “Symmetry selection rules for reaction mechanisms: Application to metal-ligand isomerizations” *Inorganic Chemistry* **1988**, *27*, 458–466
12. Johnson, B. F. G.; Rodger, A. “Principles of bonding and reactivity in transition metal cluster compounds” *Inorganica Chimica Acta* **1988**, *145*, 71–75
11. Rodger, A.; Schipper, P. E. “Symmetry selection rules for reaction mechanisms: A practical formulation for the generation of symmetry-allowed mechanisms and applications” *The Journal of Physical Chemistry* **1987**, *91*, 189–195
10. Rodger, A. “Irreducible representations for cyclic, dihedral and cubic point groups” *Australian Journal of Chemistry* **1987**, *40*, 1035–1042
9. Schipper, P. E.; Rodger, A. “Generalized selection rules for circular dichroism: A symmetry adapted perturbation model for magnetic dipole allowed transitions” *Chemical Physics* **1986**, *109*, 173–193
8. Rodger, A. “d-d Transitions of Co^{III} complexes studied by associated induced circular dichroism” *Inorganica Chimica Acta* **1986**, *122*, 25−30
7. Rodger, A.; Schipper, P. E. “Symmetry selection rules for reaction mechanisms” *Chemical Physics* **1986**, *107*, 329–342
6. Schipper, P. E.; Rodger, A. “Generalized selection rules: An augmentation procedure for the determination of symmetry invariants” *Chemical Physics* **1985**, *98*, 29–40
5. Schipper, P. E.; Rodger, A. “DICD of Co(III) complexes in sugar solutions” *Inorganica Chimica Acta* **1985**, *99*, L41−L42
4. Schipper, P. E.; Rodger, A. “Spectroscopic assignment of d-d transitions of achiral Metal complexes using circular dichroism: DICD of Co(III) complexes in sugar solutions” *Journal of the American Chemical Society* **1985**, *107*, 3459−3465

3. Rodger, A. “The kinetics of the acid catalysed hydrolysis of Sucrose” *Chemistry in Australia* **1983**, *52*, 71–71
2. Schipper, P. E.; Rodger, A. “Symmetry rules for the determination of the intercalation geometry of host / guest systems using circular dichroism: A symmetry adapted coupled-oscillator model” *Journal of the American Chemical Society* **1983**, *105*, 4541–4550
1. Ekstrom, A.; Leary, A. J.; Lindoy, L. F.; Rodger, A.; Harrison, B. A.; Tregloan, P. A. “Comparative studies of the kinetics of macrocycle dissociation from nickel (II) in the presence of excess copper ions and 1,10-phenanthroline” *Inorganic Chemistry* **1983**, *22*, 1404–1407

Patents

P5. Hannon, M.J., Rodger, A. “Supramolecular Chrial Selectors” 0222714.8 **2002**
P4. Hannon, M.J., Rodger, A. “Supramolecular Chrial Selectors” 0222628.0 **2002**
P3. Dafforn, T.R., Halsall, D.J., Rodger, A. “Linear dichroism couette flow cell design” **2002,** Patent application number 0217117.1, Filing date 24-Jul-02, Applicants University of Warwick, Cambridge University and Addenbrookes NHS trust.
P2. Dafforn, T.R., Halsall, D.J., Rodger, A. “Detection of nucleic acids” P014295GBR **2002**
P1. Dafforn, T.R., Halsall, D.J., Rodger, A. “Method for determining nucleotide mismatches” 0108329.4 **2001**

ALISON RODGER
CURRICULUM VITAE

CURRENT APOINTMENT

Professor, Department of Chemistry, University of Warwick

Director, MOAC Doctoral Training Centre, University of Warwick

CURRENT ADDRESSES

MOAC Doctoral Training Centre, University of Warwick, Coventry, CV4 7AL, UK.

PERSONAL DATA

Date of Birth: 21st November, 1959

Place of Birth: Edinburgh, Scotland

Married with two children (05.96 and 10.98)

Citizenship: Australian and UK

ACADEMIC QUALIFICATIONS

2003	*DSc*, University of Sydney
1988	*MA*, University of Oxford
1986	*PhD*, University of Sydney
1985	Diploma in Biblical studies, Moore College, Sydney
1982	*BSc* (*Hons*) (Chemistry and Mathematics), University of Sydney, Honours class I and University Medal for Theoretical Chemistry

APPOINTMENTS

2005, 2007	Visiting Professor, Stanford University, USA
2005*f*	Professor of Biophysical Chemistry, Department of Chemistry, University of Warwick
2003*f*	Director of Molecular Organisation and Assembly in Cells (MOAC) Doctoral Training Centre, University of Warwick
2003–2005	Reader, Department of Chemistry, University of Warwick
1998–2003	Senior Lecturer, Department of Chemistry, University of Warwick
1996	Visiting Professor, Stanford University, USA
1994–1998	Lecturer, Department of Chemistry, University of Warwick
1994	Visiting Professor, Stanford University, USA
1991–1994	Glasstone Fellow, University of Oxford and Inorganic Chemistry Fellow of St. Hilda's College
1992	Visiting Professor, Stanford University, USA
1987–1992	Copy-editor for Theoretica Chimica Acta (part time)
1988–1991	Unilever Fellow, St. Catherine's College, Oxford
1985–1988	Beatrice Dale Research Fellow, Newnham College, Cambridge
1985–1988	Overseas Scholar of the Royal Commission for the Exhibition of 1851

MEMBERSHIP OF PROFESSIONAL BODIES

Fellow of the Royal Australian Chemical Institute (1999)

Fellow of the Royal Society of Chemistry (2000)

RESEARCH INTERESTS: Solution phase biomacromolecule structure and function

1. Development of spectroscopic techniques, particularly linear and circular dichroism, for study of biomacromolecules.
2. Structural aspects of DNA and RNA behaviour in solution. Structure and function of DNA / ligand systems.
3. Structure and function of protein systems, particularly fibrous proteins and membrane proteins.
4. Structure and function of DNA / protein complexes.

PUBLICATIONS SUMMARY

Five books, nine refereed chapters, four theses, over 125 refereed journal articles, five patents.

THE UNIVERSITY OF WARWICK

Higher Doctorate Application Form

For Candidates for the award of Doctor of Letters, Doctor of Science, Doctor of Laws

Name	Alison Rodger	Department	Chemistry
E-mail address	a.rodger@warwick.ac.uk	Telephone	74696
Post title	Professor, Directgor MOAC Doctoral Training Centre	Start (and, if appropriate, end) date of appointment at Warwick	October 1994. Professor from Oct 2005

Please list your academic qualifications

Degree	Institution	Year of award
BSc (Hons)	University of Sydney	1982
PhD	University of Sydney	1986
MA	University of Oxford	1988
DSc	University of Sydney	2003

For which degree is the application made, ie Doctor of Letters, Doctor of Science or Doctor of Laws?

Doctor of Science

The Regulations stipulate that a candidate shall present published work(s) or a collection of published papers which shall be a record, distinguished by its high standard, of important original work by the candidate. This work shall be a major contribution to the subject.

As a guide, it is suggested that one major monograph or a minimum of three or four published papers in leading refereed journals (or a combination) should be submitted. References for each item submitted should be listed below and two copies of each must be enclosed with the application (one copy will be retained by the University).

Complete publication list for 2003-2007 to date is given below. Red papers are 4 selected review or review-type articles of AR's work.
AR's % contribution is indicated in left hand column

C9. 90%	Rodger, A.; Marrington, R.; Roper, D.; and Windsor, S. "Circular dichroism spectroscopy for the study of protein ligand interactions" in *Protein–Ligand Interactions: Methods and Protocols*, Uli Nienhaus (ed) in *Methods in Molecular Biology / Methods in Molecular Medicine* The Humana Press Inc., New Jersey, **2005**, 341–361.
123. 10%	Shelton, A.; Rodger, A.; McMillin, D. "DNA-Binding Studies of a New Di-Cationic Porphyrin. Insights into Interligand Interactions" *Biochemistry*, **2007**
122. 60%	Peberdy, J.C.; Malina, J.; Khalid, S.; Hannon, M.J.; Rodger, A. "Influence of surface shape on DNA binding of bimetallo helicates" *Journal of Inorganic Biochemistry* **2007,**
121. 33%	Bulheller, B.; Rodger, A.; Hirst, J.D. "Circular and linear dichroism of proteins" *Physical Chemistry Chemical Physics*, **2007**,
120. 15%	Small, E.; Marrington, R.; Rodger, A.; Scott, D.J.; Sloan K.; Roper, D.; Dafforn, T.R.; Addinall, S.G. "FtsZ polymer-bundling by the *Escherichia coli* ZapA orthologue, YgfE involves a conformational change in bound GTP" *Journal of Molecular Biology*, **2007**
119. 5%	Galindo, M.A.; Olea, D.; Romero, M. A.; Gómez, J.; del Castillo, P.; Hannon, M.J.; Rodger, A.; Zamora, F.; Navarro, J.A.R; "Design and Non-covalent DNA Binding of Platinum(II) Metallacalix[n]arenas" *Chemistry —A European Journal*, **2006**
118. 50%	Kerckhoffs, J.M.C.A.; Peberdy, J.C.; Meistermann, I.; Childs, L.J.; Isaac, C.J.; Pearmund, C.R.; Reudegger, V.; Alcock, N.W.; Hannon, M.J.; Rodger, A. "Enantiomeric resolution of supramolecular helicates with different surface topographies" *Dalton Transactions*, **2007**, 734–742

117. 10%	Lee, S.C.; Stoilova-Mcphie, S.; Baxter, L.; Fülöp, V.; Henderson, J.; Rodger, A.; Roper, D.; Scott, D.J.; Smith, C.; Morgan, J.A.W; "Biophysical characterisation of an insecticidal toxin component XptA1 suggest a new toxin mode of action" *Journal of Molecular Biology*, **2007**: doi:10.1016/j.jmb.2006.12.057
116. 50%	Richards, A.D.; Rodger, A. "Synthetic metallomolecules as agents for the control of DNA structure" *Chemical Society Reviews*, **2007**, *36*, 471–483
115. 15%	Duff, M.R; Tan, W.B.; Bhambhani, A.; Perrin, B.S. Jr.; Thota, J.; Rodger, A.; Kumar, C.V. "Contributions of hydroxyethyl groups to the DNA binding affinities of anthracene probes" *The Journal of Physical Chemistr,y* **2006**
114. 10%	Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. "Shape effects on the activity of synthetic major-groove binding ligands" *Journal of Molecular Graphics and Modelling*, **2006** doi:10.1016/j.jmgm.2006.07.004
113. 10%	Stootman, F.H.; Fisher, D.M.; Rodger, A.; Aldrich-Wright, J.R. "Improved curve fitting procedures to determine equilibrium binding constants" *The Analyst*, **2006**
112. 60%	Rodger, A.; Marrington, R.; Geeves, M.A.; Hicks, M.; de Alwis, L.; Halsall, D.J.; Dafforn, T.R. "Looking at long molecules in solution: what happens when they are subjected to Couette flow?" *Physical Chemistry Chemical Physics*, **2006**, *8*, 3161–3171
111. 50%	Hannon, M.J.; Green, P.S.; Fisher, D.M.; Derrick, P.J.; Beck, J.L.; Watt, S.J.; Ralph, S.F.; Shiel, M.M.; Barker, P.R.; Alcock, N.W.; Price, R.J.; Sanders, K.J.; Pither, R.; Davis, J.; Rodger, A. "An estrogen-platinum terpyridine conjugate: DNA and protein binding and cellular delivery", *Chemistry —A European Journal*, **2006**
110. 15%	Hicks, M.R.; Rodger, A.; Thomas, C.M.; Batt, S.M.; Dafforn, T.R. "Restriction enzyme kinetics monitored by UV linear dichroism" *Biochemistry* **2006**, *45*, 8912–8917
109. 50%	Marrington, R.; Seymour, M.; Rodger, A. "A new method for fibrous protein analysis illustrated by application to tubulin microtubule polymerisation and depolymerisation" *Chirality* **2006**
108. 10%	Childs, L.J.; Malina, J.; Pascu, M.; Rolfsnes, B.E.; Prieto, M.J.; Broome, M.J.; Rodger, P.M.; Sletten, E.; Moreno, V.; Rodger, A.; Hannon, M.J. "A DNA-binding cuprous metallo-supramolecular cylinder that can act as an artificial nuclease" *Chemistry —A European Journal* **2006**
107. 10%	Oleksi, A.; Blanco, A.G.; Boer, R.; Usón, I.; Aymamí, J.; Rodger, A.; Hannon, M.J.; Coll, M. "Molecular recognition of a three-way DNA junction by a metallo-supramolecular helicate", *Angewandte Chemie* **2006**, 45,1227–
106. 10%	Khalid, S.; Hannon, M.J.; Rodger, A.; Rodger, P.M. "Simulations of DNA coiling around a synthetic supramolecular cylinder that binds in the DNA major groove" *Chemistry —A European Journal* **2006**, *12*, 3493-506
105. 50%	Rajendra, J.; Damianoglou, A.; Hicks, M.; Booth, P.; Rodger, P.M.; Rodger, A. "Quantitation of protein orientation in flow-oriented unilamellar liposomes by linear dichroism", *Chemical Physics* **2006**, *326*, 210–220
104. 15%	Tan, W. B., Bhambhani, A., Duff, M. R., Rodger, A. and Kumar, C. V. (2006) Spectroscopic identification of binding modes of anthracene probes and DNA sequence recognition. *Photochemistry and Photobiology* **2006**, *82*, 20–30
103. 40%	Marrington,R.; Dafforn, T.R.; Halsall, D.J.; Hicks, M.; Rodger, A. "Validation of new microvolume Couette flow linear dichroism cells" *Analyst*, **2005**, *130*, 1608–1616
102. 70%	Khalid, S.; Rodger, P.M.; and Rodger,A. "Theoretical aspects of the enantiomeric resolution of dimetallo helicates with different surface topologies on cellulose columns" *Journal of Liquid Chromatography* **2005**, *28*, 2995–3003
101. 25%	Koutsoudaki, C.; Kresk, M.; Rodger, A. "Chemical composition and antibacterial activity of the essential oil and the gum of *Pistacia lentiscus var. chia*" *Journal of Agricultural and Food Chemistry* **2005**, *53*, 7681–7685

100. 50%	Rajendra, J.; Rodger, A. "The binding of single stranded DNA and PNA to single walled carbon nanotubes probed by flow linear dichroism" *Chemistry - A European Journal* **2005**, *11*, 4841–48
99. 5%	Addinall, S. G.; Johnson, K.A.; Dafforn, T.R.; Smith, C.; Rodger, A.; Gomez, R.P.; Sloan, K.; Blewitt, A.; Scott, D.J.; Roper, D.I "Expression, purification and crystallisation of the cell-division protein YgfE from *Escherichia Coli*" *Acta Cryst. Structural Biology and Crystallisation Communications* **2005,** *F61*, 305–307
98. 5%	Uerpmann, C.; Malina, J.; Pascu, M.; Clarkson, G.J.; Moreno, V.; Rodger, A.; Grandas, A.; Hannon, M.J. "Design and DNA binding of an extended triple-stranded metallo-supramolecular cylinder" *Chemistry - A European Journal* **2005**
97. 20%	Halsall, D.J.; Dafforn, T.R; Marrington, R.; Halligan, E.; and Rodger, A. "Linear Dichroism for quantitative Polymerase Chain Reaction applications: An intrinsic signal from DNA has the potential to be an interference-free, specific and label free probe for real-time and quantitative PCR" *IVD Technology*, **2004**, *6*, 51–60
96. 50%	Dafforn, T.R. and Rodger, A. "Unravelling the configuration of protein fibres and membrane proteins" *Current Opinion in Structural Biology*, **2004**, *14*, 541–546
95. 20%	Marrington, R.; Small, E.; Rodger, A.; Dafforn, T.R.; Addinall, S. "FtsZ fibre bundling is triggered by a calcium-induced conformational change in bound GTP" *Journal of Biological Chemistry* **2004**, *47*, 48821–48829
94. 40%	Rajendra, J.; Baxendale, M.; Dit Rap, L.G.; Rodger, A. "Flow linear dichroism to probe binding of aromatic molecules and DNA to single walled carbon nanotubes" *Journal of the American Chemical Society*, **2004**, *126*, 11182–11188
93. 15%	Ambroż, H.B.; Kemp, T.J.; Rodger, A.; and Przybytniak, G. "Ferric and ferrous ions: binding to DNA and influence on radiation-induced processes" *Radiation Physics and Chemistry*, **2004**, *71*, 1023–1030
92. 40%	Marrington, R.; Dafforn, T. R.; Halsall, D.J.; Rodger, A. "Micro volume couette flow sample orientation for absorbance and fluorescence linear dichroism" *Biophysical Journal*, **2004**, *87*, 2002–2012
91. 5%	McCabe, R.W., Rodger, A. and Taylor, A. "A study of the secondary structure of *Candida antarctica* lipase B using synchrotron radiation circular dichroism measurements" *Enzyme and Microbial Technology*, **2004**, *36*, 70–74
90. 33%	Alcock, N.W., Sanders, K.J., Rodger, A. "Potential injectable contraceptive steroids: testosterone buciclate" *Acta Crystallographica* **2004**, *E60*, o348–o349
89. 25%	Dafforn, T.R.; Rajendra, J.; Halsall, D.J.; Serpell, L.C.; Rodger, A. "Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell" *Biophysical Journal* **2004**, *86*, 404–410
88. 10%	Glover, P.B.; Ashton, P.R.; Childs, L.J.; Rodger, A.; Kercher, M., Williams, R.M., Luisa De Cola, L. and Pikramenou, Z., "Hairpin-Shaped Heterometallic Luminescent Lanthanide Complexes for DNA Intercalative Recognition" *Journal of the American Chemical Society* **2003**, *125*, 9918–9919
87. 50%	Hannon, M.J. and Rodger, A. "Groovy DNA recognition: new agents for gene regulation?" *Pharmaceutical Visions*, **2003**, 14–16
86. 20%	Miguel, M.S., Marrington, R., Rodger, P.M., Rodger, A., and Robinson, C. "An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on hydrophobicity of the environment" *European Journal of Biochemistry* **2003**, *270*, 3345–3352
85. 15%	Miles, A.J., Wien, F., Lees, J.G., Rodger, A., Janes, R.W. and Wallace, B.A. "Calibration and Standardisation of Synchrotron Radiation Circular Dichroism and Conventional Circular Dichroism Spectrophotometers" *Spectroscopy* **2003**, 17, 653–661
84. 5%	Novakova, O., Chen, H., Vrana, O., Rodger, A., Sadler, P.J., Brabec, V. "DNA interactions of monofunctional organometallic ruthenium(II) antitumor complexes in cell-free media" *Biochemistry* **2003,** *42*, 11544–11554

83. 15%	Đapic, V., Abdomerovic, V., Marrington, R., Peberdy, J.C., Rodger, A., Trent, J.O., Bates, P.J., "Biophysical and biological properties of G-quartet forming oligonucleotides" *Nucleic Acids Research* **2003**, *31*, 2097–2107
82. 20%	Grummitt, A.R., Harding, M.M., Anderberg, P.I., Rodger, A. "Carbohydrate derivatives of the antitumour alkaloid 9-hydroxyellipticine" *European Journal of Organic Chemistry* **2003**, *1*, 63–71

If any part of the work was undertaken in collaboration with other persons, a statement of the extent of the contribution by such persons must be indicated below:

The key feature of Alison Rodger's research is its multidisciplinary nature. Thus nearly all published work is undertaken in collaboration.

Please also indicate in the box below whether or not the work or any part of it has been submitted, successfully or unsuccessfully, for a degree in this or any other University.

None of the above work has been submitted for another degree. Nearly all publications prior to the above listed ones have been.

Please sign and date in the box below (or type, if submitting form electronically)

Alison Rodger

Submit
Please forward this form as an email attachment to: j.h.gardner@warwick.ac.uk, or post to: Jocasta Gardner, Graduate School, University House.

Please also attach a summary CV (ie one/two page) to your submission. Please note that the purpose of the CV is to enable your submitted works to be placed in the context of your significant publications during the relevant period. It will be sent to those considering your application, including where appropriate to the examiners, for this purpose.

Thank you for taking the time to complete this form. We will contact you in due course to advise you of progress.

The following are for use by the Graduate School and are not to be completed by the candidate

Date received in Graduate School:		Date of enrolment (ie commencement date entered on SITS):	
Date of Acknowledgement to the candidate:		Expected completion date (normally 6 months after enrolment):	
Initial consideration by Chair of Faculty Graduate Studies Committee and Head of Department:	**Recommendation:**	**Rationale for recommendation:**	

Consideration by Board of Graduate Studies	**Decision (and date):**	Consideration by the Senate:	**Decision (and date):**
Name and Address of Examiner 1		Name and Address of Examiner 2	
The reports of the Examiners, once received, must be attached to this form			
Recommendation to the Senate:			
Date of Approval of Award by the Senate		Fee paid (N/A in 2006/07)	**N/A**
Graduate School Office Notes			

REGULATIONS (attached for information)

Doctor of Letters, Doctor of Science, Doctor of Laws

(1) Candidates for these degrees must:

(a) be graduates of the University of Warwick of seven years' standing; *or*

(b) have been members of staff or full-time research workers of the University of Warwick for four years and be graduates of seven years' standing of a University recognised for this purpose by the Senate.

(2) A candidate shall present published work(s) or a collection of published papers which shall be a record, distinguished by its high standard, of important original work by the candidate. This work shall be a major contribution to the subject. Where any part of the work was done in collaboration with other persons, a statement of the extent of the contribution by such persons shall be included. The candidate shall also indicate whether or not the work or any part of it has been submitted, successfully or unsuccessfully, for a degree in this or any other University.

(3) Each application for a higher doctorate shall first be considered by the Chair of the Graduate Studies Committee of the appropriate Faculty and the Head of the appropriate Department who shall submit a joint report to the Board of Graduate Studies advising the Committee as to whether the application is worthy, prima facie, of consideration and specifying the reasons for their recommendation. The Board of Graduate Studies shall then determine whether the application is worthy, prima facie, of further consideration. After such approval, examiners for the application shall be appointed by Senate. There shall be at least two examiners, one of whom may be an internal examiner. The examiners shall submit individual reports together with a joint recommendation to the Chair of the Board of Graduate Studies for onward transmission to the Senate. Where they cannot reach a joint recommendation a further external examiner shall be appointed who shall be given access to the original reports and shall make a final recommendation. For applications by members of staff of the University, all the examiners shall be external examiners.

(4) A fee of £600 must be paid and two copies of the published work submitted with the application. If the application is declared not prima facie worthy of consideration by the Board of Graduate Studies, £400 of the fee shall be returned to the candidate.

(5) One copy of the work submitted in support of a successful application shall be retained by the University.

[Candidates should note that the fee, under Regulation (4) above, has been waived for those applying during the 2006/07 academic year]

Carbohydrate Derivatives of the Antitumour Alkaloid 9-Hydroxyellipticine

Annaleise R. Grummitt,[a] Margaret M. Harding,*[a] Pia I. Anderberg,[a] and Alison Rodger[b]

Keywords: Glycosides / Antitumour agents / Heterocycles / DNA / Drug design

The synthesis of L-arabinosyl derivatives of 2-naphthol and the quaternised ellipticines Celiptium (**2**) and Ellipravin (**3**) is reported. Naphth-2-yl 2′,3′,4′-tri-*O*-acetyl-α-L-arabinopyranoside was prepared under Königs–Knorr and Mitsunobu conditions in nonpolar aprotic solvents and using 2,3,4-tri-*O*-acetyl-L-arabinopyranosyl fluoride as the glycosyl donor. These conditions were not applicable to the corresponding glycosidation reactions with the quaternised ellipticines **2** and **3** which are soluble only in polar solvents. Formation of the 9-(α-L-arabinopyranosyl)ellipticine derivatives **13** and **14** was achieved by using 2,3,4-tri-*O*-acetyl-L-arabinopyranosyl bromide in the presence of sodium methoxide in methanol. Improved yields were obtained under the same conditions by incorporation of an ether linker between the sugar and ellipticine to give derivatives **15** and **16**. The glycolate esters **17** and **18**, which were prepared using 2-(2′,3′,4′-tri-*O*-acetyl-α-L-arabinopyranosyl)glycolic acid, undergo hydrolysis suggesting that these compounds could function as prodrugs in vivo. Linear dichroism studies of the interaction of Celiptium (**2**) and the stable L-arabinosyl ellipticine derivatives **3**, **15** and **16** with calf thymus DNA are consistent with intercalation of the ellipticine chromophore, positioning the sugars at the 2- and 9-positions in the major and minor grooves of DNA.

Introduction

Ellipticine and 9-methoxyellipticine are plant alkaloids isolated from the leaves of the plant *Ochrosia elliptica*, of the *Apocynaceae* family.[1] While both compounds exhibit high in vitro anticancer activity against a number of experimental tumours,[2] poor aqueous solubility has precluded clinical trials with these compounds. Research to address this limitation identified hydroxylation at C-9, to give 9-hydroxyellipticine (**1**) (Figure 1, part a),[3] and quaternisation at N-2,[4] as key structural modifications that led to improved aqueous solubility, increased antitumour activity and altered tissue distribution. As a result of these studies, preclinical and clinical evaluations of a number of derivatives of 9-hydroxyellipticine have been carried out, with 9-hydroxy-*N*-methylellipticinium acetate (**2**) (Celiptium®, Figure 1, part a) having entered the market for the treatment of advanced breast cancer.[5,6]

9-Hydroxyellipticine (**1**) and derivatives exert their antitumour activity via a multi-faceted mechanism that includes DNA intercalation, the generation of DNA strand breaks and interference with the action of topoisomerase II.[7–10] DNA intercalation is believed to be a key interaction that is directly related to anticancer activity. However, in contrast to many other established DNA intercalators, there are no X-ray crystal structures or NMR solution structures of 9-hydroxyellipticine (**1**) or derivatives with DNA. Two recent studies have provided important information about the nature of the drug/DNA complex.[11,12] Spectroscopic studies[12] and theoretical calculations[11] have shown that the chromophore of **1** is oriented perpendicular to the helix axis, with the 9-OH group protruding into the minor groove and the pyridine nitrogen atom lying in the major groove.[11,12] The placement of these two functional groups in the DNA grooves allows for their use as chemical handles for the attachment of other DNA-targeting groups,

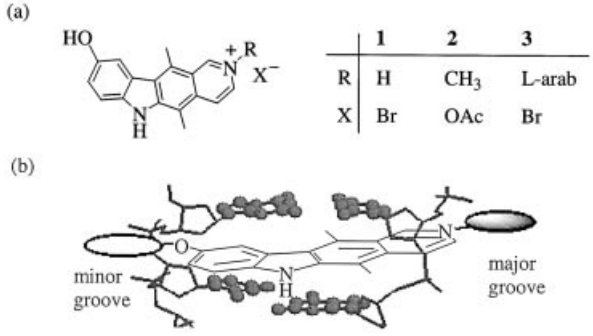


Figure 1. (a) 9-Hydroxyellipticine (**1**) and important clinical derivatives Celiptium® (**2**) and Ellipravin (**3**), where L-arab = α-L-arabinopyranosyl and (b) schematic drawing showing the relative orientation of the proposed major and minor groove binding groups at the 2- and 9-positions of ellipticines

[a] School of Chemistry, University of Sydney
N.S.W. 2006, Australia
Fax: (internat.) + 61-2/9351-6650
E-mail: harding@chem.usyd.edu.au
[b] Department of Chemistry, University of Warwick
Coventry CV4 7AL, UK

including groove binders, at these positions. Whilst a range of 2- and 9-substituted derivatives of **1** have been reported,[13] almost all derivatives contain simple alkyl and related substituents. One exception is the reported synthesis and in vitro screening of 40 *N*-glycosides of **1**.[14,15] Sugars were chosen for attachment to the ellipticine chromophore due to their prevalence in many antitumour drugs, and in order to improve aqueous solubility. 2-(α-L-Arabinopyranosyl)-9-hydroxyellipticinium bromide (**3**) (Ellipravin, Figure 1, part a) was selected from 40 glycosides for preclinical evaluation. Whilst no DNA-binding studies on **3** have been reported, modelling studies with oligonucleotides suggest that the sugar is located in the major groove.[16]

The establishment of the binding orientation of 9-hydroxyellipticine (**1**) with DNA provides the basis for the rational design of DNA threaders in which major and minor groove binding motifs are tethered to the ellipticine core (Figure 1, part b).[13] Carbohydrates are an attractive choice for the development of threaders of this type, due to the importance of sugars as groove binders in naturally occurring DNA-binding antitumour antibiotics,[17−19] and due to the significant improvements in aqueous solubility generally conferred by carbohydrates.[20] In addition, glycosylated phenols have been widely used in anticancer drug design.[21] While the exact mechanism and the mode of DNA binding of Ellipravin (**3**) are unknown, the high clinical activity of **3** has demonstrated that addition of carbohydrates to **1** has the potential to lead to novel antitumour derivatives. In this paper, we report the synthesis of six new carbohydrate derivatives of Celiptium (**2**) and Ellipravin (**3**) and preliminary DNA binding studies that confirm that these derivatives retain the ability to bind to DNA by intercalation. Of particular interest is the synthesis of a new ellipticine derivative containing two sugars, that is proposed to act as a DNA threading agent.

Results and Discussion

L-Arabinose was selected as the carbohydrate for synthetic studies, due to the demonstrated clinical activity of Ellipravin (**3**). Attachment of L-arabinose to the 9-hydroxy group of Celiptium (**2**) and Ellipravin (**3**)[15] by direct *O*-aryl glycosidation, and by the incorporation of both a short ester and a short ether linker between the sugar and the 9-position, were investigated.

Preparation of *O*-Aryl Glycosides

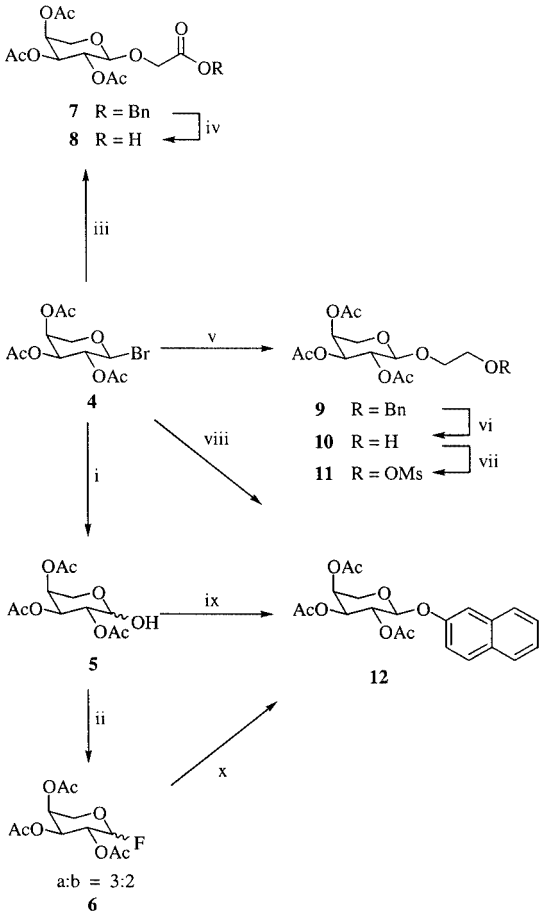
Given the limited availability of ellipticines **2** and **3**, the lack of solubility of these quaternised compounds in aprotic solvents that are generally favoured for glycosidations,[22] and the susceptibility of 9-hydroxyellipticines to oxidation, particularly under basic conditions,[8,10,13,23] glycosidation was initially investigated using 2-naphthol as a model system. The naphthyl *O*-glycoside **12** was prepared under Königs−Knorr conditions in low yield (9%) from the glycosyl bromide **4**[24] and 2-naphthol in the presence of silver carbonate (dichloromethane/diethyl ether, 1:1). Variation of the reaction conditions including the use of different solvents (toluene, acetone), the number of equivalents of reagents (1.2−5 equiv. of silver carbonate, 0.4−1.5 equiv. of sugar **4**) and the use of both silver carbonate and silver zeolite as insoluble promoters failed to significantly improve the yield of **12**. Celiptium (**2**) is insoluble in dichloromethane/diethyl ether, the most successful solvent system identified for the formation of the naphthyl-*O*-glycoside **12**. The reaction was nevertheless attempted with Celiptium (**2**) in dimethylformamide (DMF). However, there was no evidence for formation of the corresponding *O*-aryl glycoside using this solvent.

The low yields often observed in the synthesis of aryl glycosides are generally attributed to the poor nucleophilicity of the phenol functional group compared to alcohols in glycosidation reactions. This is a major contributing factor to the low observed yield in the synthesis of **12** using Königs−Knorr conditions. This limitation has been addressed recently by the application of the Mitsunobu protocol to the preparation of *O*-aryl glycosides of a number of naturally occurring DNA-binding antitumour antibiotics.[25] Using this procedure, the generation of the phenoxide ion under conditions that favoured an S_N2 pathway provided a highly efficient route for the stereoselective synthesis of aryl 2-deoxy-β-D-glycosides.

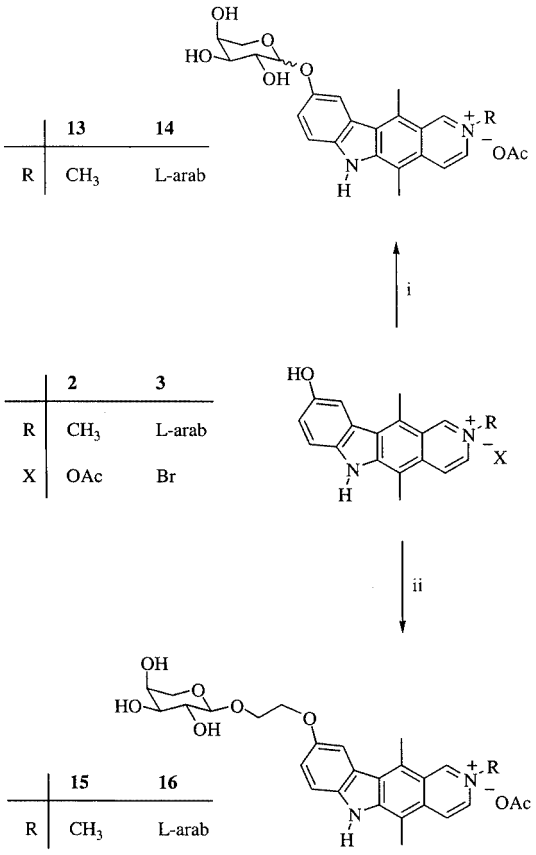
The Mitsunobu conditions[25] were first applied to the synthesis of the naphthyl *O*-glycoside **12** (Scheme 1). Treatment of hemiacetal **5**,[26] prepared by hydrolysis of the glycosyl bromide **4**, with 2-naphthol (diethyl azodicarboxylate, triphenylphosphane, 0 °C, tetrahydrofuran) afforded the desired aryl glycoside **12** in 39% yield. While this yield was higher than the yield of **12** prepared under Königs−Knorr conditions (9%), the reaction failed in DMF, the solvent of choice for the corresponding reaction with ellipticines **2** and **3**.

The third approach to form the *O*-aryl glycoside **12** involved the use of the glycosyl fluoride **6**. Glycosyl fluorides have attracted significant attention as they are readily prepared, stable to chromatography, and are excellent glycosyl donors in the presence of a range of activating agents.[22] The required glycosyl fluoride **6** was prepared as a mixture of anomers (α/β = 3:2) by treatment of hemiacetal **5** with (diethylamino)sulfur trifluoride at low temperatures (Scheme 1). Treatment of the fluoride **6** with boron trifluoride−diethyl ether followed by 2-naphthol afforded the *O*-aryl glycoside **12** in 10% estimated yield from analysis of the crude product. Due to the poor yield, this method was not investigated further.

The use of other glycosyl donors (trifluoroacetimidates,[27] thioglycosides[28]) was not investigated as glycosidations with these systems are not favoured in DMF or polar protic solvents, in which the quaternised ellipticines **2** and **3** are soluble. Similarly, initial formation of the *O*-aryl glycoside from the free-base 9-hydroxyelliptine (**1**) using the Schmidt procedure,[27] followed by *N*-glycosidation, is also problematic as **1** is also only soluble in DMF and polar protic solvents. Whilst the use of nucleophilic solvents is highly undesirable in glycosidation reactions, the single ex-



Scheme 1. (i) H_2O/acetone, Ag_2CO_3; (ii) (diethylamino)sulfur trifluoride, dichloromethane, −30 to 25 °C; (iii) $HOCH_2COOBn$, Ag zeolite, toluene, MS (4 Å); (iv) H_2, Pd/C, ethanol; (v) $HOCH_2$-CH_2OBn, Ag zeolite, toluene, MS (4 Å); (vi) H_2, Pd/C, ethanol; (vii) MsCl, NEt_3, tetrahydrofuran, 0 °C; (viii) $AgCO_3$, 2-naphthol, dichloromethane/diethyl ether (1:1); (ix) triphenylphosphane, diethyl azodicarboxylate, tetrahydrofuran, 2-naphthol; (x) boron trifluoride−diethyl ether, acetonitrile, 2-naphthol



Scheme 2. (i) **4**, sodium methoxide, methanol, 0 °C; (ii) **11**, sodium methoxide, methanol, 0 °C, where L-arab = α-L-arabinopyranosyl

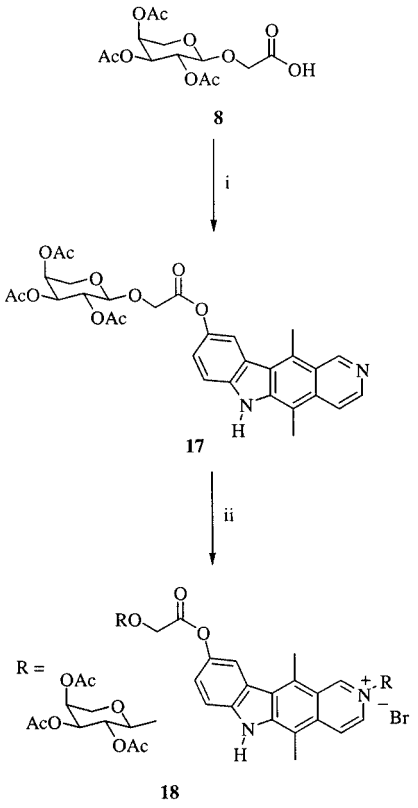
ample in the literature of *O*-glycosides of 9-hydroxyellipticine involved a low-yielding (20%) synthesis, which was accomplished by treatment of Celiptium (**2**) with a glycosyl bromide in sodium methoxide/methanol.[29] These conditions were originally reported for the *O*-glycosidation of 5-hydroxyindole[30] and were subsequently applied to the glycosidation of **2** to give *O*-aryl glycosides. Large excesses of the glycosyl bromides and short reaction times (< 5 min) were reported, presumably due to competing nucleophilic attack of the base and solvent on the sugar. Due to this literature precedent and the good solubility of Celiptium (**2**) and Ellipravin (**3**) in alcohols and water, these conditions were investigated for the synthesis of arabinosyl derivatives of 9-hydroxyellipticine.

The desired ellipticine *O*-aryl glycosides **13** and **14** were successfully prepared by treatment of Celiptium (**2**) and ellipravin (**3**), respectively, with sodium methoxide/methanol at 0 °C for 15 min in the presence of 5 equiv. of arabinosyl bromide **4** (Scheme 2). Rigorous exclusion of oxygen was required in all reactions in order to minimise oxidative degradation of ellipticine under these strongly basic conditions. In agreement with the literature,[29] a large excess of the glycosyl bromide **4**, short reaction times and low temperatures (0 °C) were essential in order to obtain the *O*-aryl glycosides **13** and **14**, which were identified by analysis of the NMR spectra and mass spectra of the crude products. However, insufficient material was obtained to allow these products to be fully characterised.

Preparation of Ester-Linked Glycosides

Given the difficulties in formation of the *O*-aryl glycosides outlined in Scheme 1, and the low overall yields of **13** and **14**, the synthesis of ester-linked derivatives of 9-hydroxyellipticine was investigated. The ester linkage was chosen as syntheses of numerous esters of 9-hydroxyellipticine (**1**) have been reported by reaction of **1** with a range of carboxylic acids in the presence of *N*,*N*′-dicyclohexylcarbodiimide and 1-hydroxybenzotriazole in DMF.[31] Hence, the carboxylic acid functional group was introduced into L-arabinose as shown in Scheme 1.

Reaction of benzyl glycolate with bromide **4** in the presence of silver zeolite afforded the protected glycoside **7** which was subjected to hydrogenolysis to give the carboxylic acid **8** in reasonable yield (Scheme 1). The acid **8** was treated with *N*,*N*′-dicyclohexylcarbodiimide and 1-hydroxybenzotriazole in DMF followed by 9-hydroxyellipticine (**1**), in the presence of a catalytic amount of (dimethylamino)pyridine to give the desired ester **17** in 17% isolated



Scheme 3. (i) (a) *N*,*N*′-dicyclohexylcarbodiimide, 1-hydroxybenzotriazole, DMF, (b) **1**, (dimethylamino)pyridine or (a) thionyl chloride, (b) **1**, $CHCl_3$, NEt_3;(ii) **4**, $CdCO_3$, nitromethane

yield (Scheme 3). A slightly improved yield (23%) was obtained by initial conversion of the acid **8** to the corresponding acid chloride, which was immediately treated with 9-hydroxyellipticine in chloroform/triethylamine (3:1) to give the protected *O*-glycoside **17**. The isolated yield of ester **17** was lower than expected from analysis of the crude product by NMR spectroscopy, as significant hydrolysis of the ester to regenerate 9-hydroxyellipticine (**1**) occurred during purification by flash chromatography.

The protected *O*-glycoside **17** was treated with glycosyl bromide **4** under the literature conditions used for the preparation of **3**.[15] The crude product contained the peracetylated bis(glycoside) **18**, as evidenced by electrospray mass spectrometry (m/z = 837 [M^+]) and analysis of the product by NMR spectroscopy, but by-products resulting from cleavage of the glycolate ester were also observed. Analytically pure material could not be obtained, as attempted purification of **18** by HPLC in aqueous ammonium acetate and methanol resulted in quantitative hydrolysis of the ester bond. Attempted deacetylation of the *O*-glycoside **17** and the bis(glycoside) **18** by treatment with ammonia in methanol at low temperatures (−15 °C) also resulted in cleavage of the glycolic ester bond. This susceptibility of the glycolate ester bond to hydrolysis is presumed to arise from the increased electrophilicity of the carbonyl group due to the inductive, electron-withdrawing effect of the ether substituent on the α-carbon atom. Thus, while the protected esters **17** and **18** were prepared, the poor stability of the ester linker suggested that these compounds would hydrolyse in vivo and hence further chemistry on these compounds was not investigated.

Preparation of Ether-Linked Glycosides

Replacement of the ester linker with the hydrolytically stable ether functional group was achieved as shown in Scheme 2. Treatment of the arabinosyl bromide **4** with 2-(benzyloxy)ethanol afforded the protected arabinose derivative **9**, which was converted into the mesylate **11** under standard conditions via the alcohol **10** (Scheme 1). Initial attempts to synthesise **15** and **16** using mild bases (Cs_2CO_3, K_2CO_3) in DMF were unsuccessful. The ellipticine *O*-glycoside **15** and bis(glycoside) **16** were both successfully prepared by treatment of the 9-hydroxyellipticines **2** and **3**, respectively, with sodium methoxide in methanol at 0 °C for 15 min in the presence of 5 equiv. of mesylate **11** (Scheme 2).

The reaction conditions were optimised by monitoring the reaction of Celiptium (**2**) with mesylate **11** by electrospray mass spectrometry, as the positively charged ellipticines in the reaction mixture give strong positive molecular ions, allowing aliquots of the reaction to be rapidly analysed at short time intervals. The reaction time of 15 min, as well as the presence of at least 5 equiv. of the mesylate **4**, resulted in maximum formation of the glycoside **15**. With less than 5 equiv. of mesylate, unchanged and degraded starting materials were recovered, while prolonged reaction times (30 min to 2 h) resulted in the formation of a higher molecular weight peak by mass spectral analysis, corresponding to a bis(sugar) adduct of Celiptium (**2**) at the expense of the desired *O*-glycoside **15**; this by-product was not isolated or characterised. While the isolated yields of analytically pure **15** and **16** after HPLC were very low (< 5%), the small scale of the reactions and the purification of charged salts most likely contribute to this figure; the mass spectral analysis of the reaction with time suggest that the glycosides are formed in approximately 20% yield. Hence it is possible that on a larger scale an improved isolated yield could be obtained.

DNA-Binding Studies

The interactions of Celiptium (**2**), Ellipravin (**3**), *O*-glycoside **15** and bis(glycoside) **16** with calf thymus DNA (ct-DNA) were studied by flow linear dichroism (LD) spectroscopy. In this technique, the DNA is oriented with respect to the incident (linearly polarised) radiation allowing the interaction of the oriented DNA with a ligand that has absorption bands in the UV/Vis region to be monitored. Any ligand that becomes oriented as a result of becoming associated with the DNA will give rise to an LD signal in the chromophore absorption band of the ligand with an intensity that is usually proportional to the amount of bound ligand.[32,33]

LD has been used previously to study the interaction of the parent 9-hydroxyellipticine **1** with DNA.[12] Titration of a solution of **1** into a solution of ct-DNA at constant DNA

concentration showed that the binding mode is dependent on the ratio of the drug to the DNA base-pairs.[12] Initial titration studies with Celiptium (**2**) using LD and CD showed similar trends to those reported for 9-hydroxyellipticine (**1**) (data not shown). Hence, LD studies with ellipticines **2**, **3**, **15** and **16** were carried out at a drug/DNA ratio of 1:10. Under these conditions, all the potential intercalation sites are not saturated and a single binding mode is present.

Figure 2 shows the UV and LD spectra obtained on addition of each of the ellipticines **2**, **3**, **15** and **16** to ct-DNA. The UV spectra of the ellipticines (curve i, upper spectra) overlap with the DNA absorptions below 300 nm (data not shown). Upon addition of Celiptium (**2**) to DNA (Figure 2, part a, curve ii), the Celiptium (**2**) absorption maxima at 326, 371 and 465 nm shift to longer wavelengths, consistent with formation of a Celiptium/DNA complex. Similar trends are observed in the UV spectra of ellipticines **3**, **15** and **16** (Figure 2, parts b−d) in the presence of ct-DNA.

The lower curves in Figure 2 show the corresponding LD spectra of the ellipticine plus ct-DNA solutions (curve iii). In each case, the presence of an LD signal at wavelengths corresponding to the maxima in the absorbance spectrum of the ellipticine plus DNA solution is immediate evidence that all four ellipticines interact with DNA and become oriented. In addition, the negative LD signal for the aromatic transitions of the ellipticines above 300 nm, where there is no overlap with the DNA absorptions, necessitates that the ellipticine chromophore is more perpendicular than parallel to the DNA helical axis, consistent with intercalation between the base pairs. For comparison, groove binders typically exhibit strong variations of the LD signal with wavelength, giving both positive and negative signals.[32,33] Small increases in the LD signals of the four ellipticine/DNA complexes at 260 nm, compared with the LD of the DNA alone (data not shown), are also consistent with intercalation, which occurs due to lengthening or stiffening of the DNA about the intercalation site.

The LD^r spectra (the difference between the normal absorption spectrum and the linear dichroism spectrum), shown in the lower part of Figure 2 (curve iv) give further information about the binding mode. The relatively flat and negative LD^r spectra confirm the orientation of the ellipticine chromophores perpendicular to the long axis of the DNA. In addition, the increase in the LD^r signal at wavelengths corresponding to the ellipticines (i.e., > 300 nm in the region where there is no overlap with the DNA absorptions) in comparison to those due to the DNA, confirm lengthening and stiffening of the DNA, consistent with the intercalation of the compound. The strong similarity between the spectra in Figure 2, parts a−d indicates a similar mode of binding for all compounds.

From these data alone, the relative orientation of the ellipticine chromophores with respect to the DNA bases (i.e., whether the pyridine nitrogen atom projects into the major or the minor groove) cannot be determined. The characteristics of the DNA plus Celiptium (**2**) spectrum (Figure 2, part a) are almost identical to the published data for DNA plus 9-hydroxyellipticine **1**.[12] This is not surprising since the addition of a single methyl group at the 2-position, which converts **1** into **2**, would not be expected to alter the binding orientation of the compound with DNA. Further-

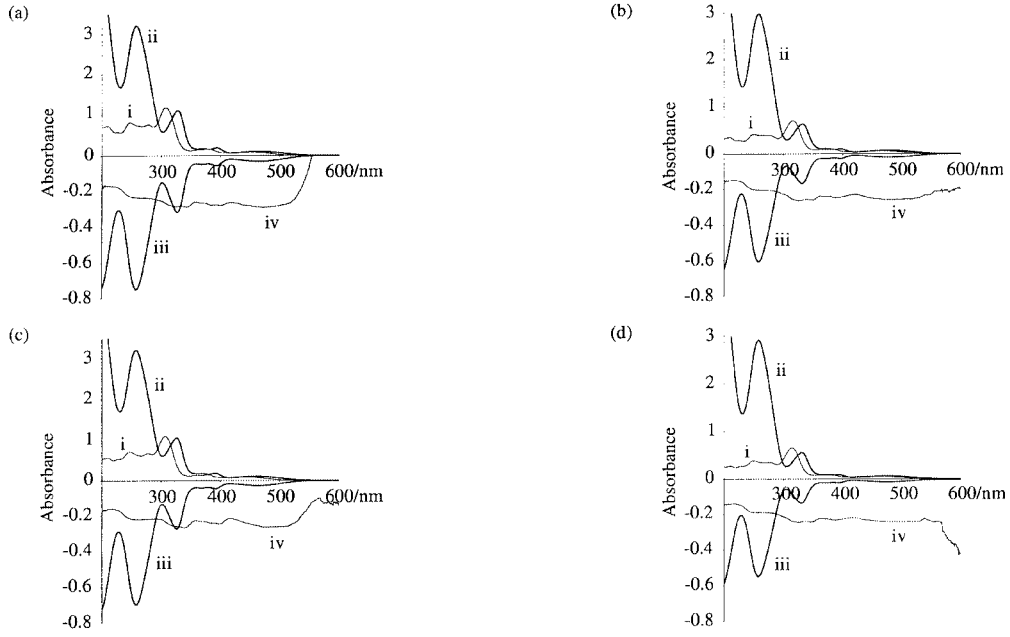


Figure 2. UV and LD spectra showing the binding of ellipticines **2** (a) **3** (b) **15** (c) and **16** (d) to calf thymus DNA at pH = 7; the UV spectra show the ellipticines (40 μM) (curve i) and the ellipticines (50 μM) in the presence of ct-DNA (500 μM) (curve ii); shown below are the corresponding LD spectra (curve iii) and LD^r spectra (curve iv) of the ellipticine plus DNA solutions; the LD spectra are multiplied by 10 to correct for different cell path lengths

more, the steric bulk of the additional glycosides in compounds **3**, **15** and **16** would be expected to strongly disfavour any orientation that does not allow these groups to project out into the grooves of the DNA.

The bis(glycoside) **16** is of particular interest as a new class of DNA-threading agent. Intercalation of the chromophore is only possible if the molecule threads through the DNA, positioning one sugar in the major groove and one sugar in the minor groove (see Figure 1, part b), with the potential to provide multiple sites of interaction between the DNA and the threader. Further structural characterisation of this complex, including the measurement of equilibrium association constants and dissociation rate constants would be informative in establishing whether the two sugars significantly increase the residence time of **16** on DNA compared with Celiptium (**2**).

Conclusion

The *O*-aryl ester- and ether-linked arabinosyl derivatives of Celiptium and ellipravin **13**−**18** represent a new class of derivatives of 9-hydroxyellipticine. While the limited stability of the ester derivatives **17** and **18** suggests that these compounds could function as pro-drugs in vivo, the stable bis(glycoside) **16** is of particular interest as a DNA threading agent; DNA intercalation of the ellipticine chromophore of **16** positions one sugar in the major groove and one sugar in the minor groove. While the impact on biological activity of the sugars at the 9-position in **13**−**16** cannot be predicted, the absence of the phenol functional group at the 9-position will prevent the formation of reactive quinone imines; formation and subsequent alkylation of these intermediates is a major pathway related to the cytoxicity of 9-hydroxyellipticines including **1** and **2**.[8,10] This change in chemical reactivity, as well as the effect of the new sugars on DNA binding, topoisomerase II poison activity and inhibition of p53 phosphorylation, should provide valuable information on the in vivo mechanism of action of the ellipticines and may identify new lead compounds with interesting biological activity profiles.

Experimental Section

General Remarks: Reactions requiring anhydrous conditions were performed under nitrogen in oven-dried glassware using freshly distilled and purified solvents. Flash chromatography was performed on Merck silica gel (type 9385, 230−400 mesh). HPLC was performed using a Waters 600E Multisolvent Delivery System, a Rheodyne 7725i injector and a Waters 486 tuneable absorbance detector (254 nm analytical; 350 nm preparative), and Alltech Alltima C18 columns: analytical 5 micron column (250 mm × 4.6 mm i.d.), flow rate 0.8 mL min^{-1}; preparative 10 micron column (300 mm × 22 mm i.d.), flow rate 7 mL min^{-1}. The solvents used were 50 mM ammonium acetate buffer (solvent A) and methanol (solvent B). Melting points were determined using a Reichert heating stage and are uncorrected. Optical rotations were recorded at 25 °C with a Perkin−Elmer 241 polarimeter. $[\alpha]_D$ values are given in 10^{-1}deg·cm^2·g^{-1}. Microanalyses were performed by the Microanalytical Unit at the University of Otago, New Zealand. Ultraviolet spectra were recorded with a Cary 4E UV/Vis/NIR spectrophotometer at 20 °C. Infrared spectra were recorded as thin films with a Perkin−Elmer 1600 Fourier transform spectrophotometer. ^{1}H NMR spectra were recorded with a Bruker AC200 NMR spectrometer or a Bruker WM AMX 400 and referenced to the residual solvent peak. EI mass spectra were recorded with an AEI MS902 spectrometer, CI mass spectra were recorded with a ThermoQuest PolarisQ Ion Trap mass spectrometer and ES mass spectra were recorded with an LCQ Ion Trap mass spectrometer, using capillary voltages of +10 V (positive ion) and −10 V (negative ion). High-resolution mass spectra were recorded as follows. EI: 70 eV with a source temperature of 200 °C and a 5.3 kV accelerating voltage, using perfluorokerosene as the reference mass; ES: Kratos Analytical Concept ISQ mass spectrometer to a resolution of 7000 (10% valley definition); LSIMS: recorded using a 10 kV caesium ion primary beam, with *m*-nitrobenzoic acid as a liquid matrix and internal mass reference. 9-Acetoxyellipticine was a gift from Professor M. Ishiguro, Suntori Institute for Bioorganic Chemistry, Osaka and Celiptium (**3**) was provided by Sanofi Chimie, Sisteron.

2,3,4-Tri-*O*-acetyl-L-arabinose (5): Silver carbonate (54.5 mg, 0.198 mmol) was added to a stirred solution of bromide **4**[24] (44.0 mg, 0.130 mmol) in acetone (0.80 mL) and water (9.0 μL, 0.50 mmol) at 0 °C with minimum exposure to light. After 4 h, the insoluble silver salts were removed by filtration (0.45 μm Millipore filter) and the solvent removed to give hemiacetal **5**[26] (35.3 mg, 98%) as a cream-coloured oily solid; α/β = 1:2. ^{1}H NMR (400 MHz, $CDCl_3$, 27 °C): δ = 5.47 (d, $^3J_{1\beta,2}$ = 3.4 Hz, 1-H_β), 4.61 (d, $^3J_{1\alpha,2}$ = 7.5 Hz,1-H_α) ppm.

2,3,4-Tri-*O*-acetyl-L-arabinopyranosyl Fluoride (6): Hemiacetal **5** (390 mg, 1.4 mmol) was dissolved in dichloromethane (20 mL) and chilled to −30 °C with stirring under nitrogen. (Diethylamino)sulfur trifluoride (250 μL, 1.9 mmol) was added rapidly and the solution allowed to warm to room temperature. After 1 h, the solution was chilled to −30 °C and methanol (3 mL) added. The solvent was removed, the residue taken up in chloroform (60 mL), washed with water (5 × 40 mL), dried with anhydrous sodium sulfate and the solvent removed to give the crude product (370 mg). Purification by flash chromatography (dichloromethane) afforded sugar **6** (338 mg, 86%) as a clear oil; α/β = 3:2. $C_{11}H_{15}FO_7$ (278.2): calcd. C 47.48, H 5.43; found C 47.67, H 5.48. IR: $\tilde{\nu}_{max}$ = 1748 cm^{-1} (CO). ^{1}H NMR (400 MHz, $CDCl_3$, 27 °C): δ = 5.78 (dd, $^3J_{1\beta,2\beta}$ = 2.7, $^3J_{1\beta,F}$ = 54.5 Hz, 1-Hβ), 5.42 (m, 4-H), 5.37 (dd, $^3J_{3\beta,4\beta}$ = 3.4, $^3J_{3\beta,2\beta}$ = 10.7 Hz, 3-H_β), 5.30−5.22 (m), 5.16 (ddd, $^3J_{2\beta,1\beta}$ = 2.7, $^3J_{2\beta,3\beta}$ = 10.7, $^3J_{2\beta,F}$ = 29.8 Hz, 2-H_β), 4.12 (m), 3.88 (dd, $^3J_{5e,4}$ = 2.0, $^3J_{5e,5a}$ = 13.3 Hz, 5-H_e), 3.77 (ddd, $^3J_{5a,4}$ = 1.0, $^3J_{5a,5e}$ = 11.5 Hz, 5-H_a) 2.17−2.03 (m, 9 H, CH_3CO's) ppm. ^{19}F NMR (376 MHz, $CDCl_3$, 27 °C): δ = −133.7 (d, *J* = 49 Hz), −149.2 (dd, *J* = 23, *J* = 53 Hz) ppm. MS (EI): *m/z* (%) = 278 (3) [M^+], 259.0816 [M − F˙] {$C_{11}H_{15}O_7$ requires 259.0818}, 115 (100), 216 (5) [M − F − Ac].

2-(2′,3′,4′-Tri-*O*-acetyl-α-L-arabinopyranosyl)glycolic Acid (8): Silver zeolite (2 g) and activated 4-Å molecular sieves were added to a solution of benzyl glycolate (153 mg, 0.92 mmol) in toluene (12 mL) and the mixture stirred in the dark for 5 min at room temperature. A solution of bromide **4** (462 mg, 1.6 mmol) in toluene (30 mL) was added and the reaction mixture stirred at room temperature. Further portions of silver zeolite (1 g) were added after 11.5 h and 23 h. After 39 h, the reaction mixture was diluted with dichloromethane, filtered through silica, and the solvent was re-

moved to give the crude product (530 mg). Purification by flash chromatography (0.8 → 3% acetone in dichloromethane) afforded glycolate **7** as a white solid (342 mg, 87%). $[\alpha]_D^{20} = -14$ (c = 4.2, $CHCl_3$). ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 7.35 (s, 5 H, Ph), 5.3−5.2 (m, 2 H, 2′-H and 4′-H), 5.18 (s, 2 H, 2 × 2-H), 5.05 (dd, $^3J_{3',4'}$ = 3.5, $^3J_{3',2'}$ = 9.3 Hz, 1 H, 3′-H), 4.58 (d, $^3J_{1',2'}$ = 6.6 Hz, 1 H, 1′-H_a), 4.32 (s, 2 H, PhCH_2), 4.02 (dd, $^3J_{5e',5a'}$ = 13, $^3J_{5e',4'}$ = 3.5 Hz, 1 H, 5′-H_e), 3.60 (dd, $^3J_{5a',5e'}$ = 12.8, $^3J_{5a',4'}$ = 1.8 Hz, 1 H, 5′-H_a), 2.12 (s, 3 H, CH_3CO), 2.06 (s, 3 H, CH_3CO), 2.02 (s, 3 H, CH_3CO) ppm. The glycolate **7** (328 mg, 770 μmol) was dissolved in ethanol (35 mL), Pd/C (10%, 320 mg) added and the mixture stirred under hydrogen for 15.5 h. The reaction mixture was filtered through Celite and the solvent removed to give the crude product (220 mg) which was dissolved in diethyl ether and extracted into saturated $NaHCO_3$. The combined aqueous extracts were acidified to pH = 3 with hydrochloric acid (10 M) and extracted into dichloromethane. The combined organic extracts were dried with anhydrous sodium sulfate and the solvent was removed to give sugar **8** (190 mg, 72%). $[\alpha]_D^{20} = -7.0$ (c = 6.3, $CHCl_3$). ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 5.30−5.20 (m, 2 H, 2′-H and 4′-H), 5.09 (dd, $^3J_{3',4'}$ = 3.5, $^3J_{3',2'}$ = 9.3 Hz, 1 H, 3′-H), 4.57 (d, $^3J_{1',2'}$ = 6.5 Hz, 1 H, 1′-H), 4.34 (s, 2 H, 2 × 2-H), 4.06 (dd, $^3J_{5e',4'}$ = 3.3, $^3J_{5e',5a'}$ = 13.1 Hz, 1 H, 5′-H_e), 3.68 (d, $^3J_{5a',5e'}$ = 13.2 Hz, 1 H, 5′-H_a), 2.15 (s, 3 H, CH_3CO), 2.11 (s, 3 H, CH_3CO) and 2.04 (s, 3 H, CH_3CO) ppm. MS (ES^-): m/z (%) = 668 (100) [M_2], 334 (72) [M]. MS (CI^-): m/z (%) = 334 (12) [M], 333 (100) [M − H], 291 (40) [M − Ac]. LSIMS: m/z = 335.0973 [M + H^+] {$C_{13}H_{19}O_{10}$ requires 335.0978}.

2′-(Benzyloxy)ethyl 2,3,4-Tri-*O*-acetyl-α-L-arabinopyranoside (9): Silver zeolite (5.14 g) and molecular sieves (4 Å) were added to a solution of 2-(benzyloxy)ethanol (0.79 mL, 5.6 mmol) in toluene (50 mL). A solution of bromide **4** (3.20 g, 9.4 mmol) in distilled toluene (300 mL) was added and the mixture stirred at room temperature under nitrogen for 5 d, filtered through Celite and the solvent removed to give the crude product (3.46 g). Purification by flash chromatography (25 → 30% EtOAc in hexane) gave sugar **9** (2.05 g, 90%) as a clear oil. $[\alpha]_D^{20} = +5.8$ (c = 5.0, $CHCl_3$): IR: $\tilde{\nu}_{max}$ = 1742 (CO) cm^{-1}. ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 7.28−7.21 (m, 5 H, Ph), 5.19−5.11 (m, 2 H, 2-H and 4-H), 4.97 (dd, ^3J = 3.5, 9.4 Hz, 1 H, 3-H), 4.48 (s, 2 H, PhCH_2), 4.46 (d, ^3J = 6.9 Hz, 1 H, 1-H), 4.00−3.86 (m, 3 H, 2 × 2′-H and 5-H_e), 3.73−3.52 (m, 3 H, 2 × 1′-H and 5-H_a), 2.06 (s, 3 H, CH_3CO), 1.95 (6 H, 2 × CH_3CO) ppm. MS (EI): m/z (%) = 410 (3) [M], 367.1389 [M − Ac] {$C_{18}H_{23}O_8$ requires 367.1393}, 91 (100).

2′-Hydroxyethyl 2,3,4-Tri-*O*-acetyl-α-L-arabinopyranoside (10): Sugar **9** (1.54 g, 3.75 mmol) was dissolved in ethanol (150 mL), Pd/C (10%, 1.57 g) was added, and the mixture stirred under hydrogen for 22 h. The mixture was filtered through Celite and the solvent removed to give the crude product (1.23 g), which was purified by flash chromatography (0 → 2% methanol in dichloromethane) to give sugar **10** (1.15 g, 96%) as a cream-colored, crystalline solid; m.p. 90.5−91.5 °C. $[\alpha]_D^{20} = +53$ (c = 1.9, $CHCl_3$). $C_{13}H_{20}O_9$ (320.3): calcd. C 48.75, H 6.29; found C 49.03, H 6.44%. IR: $\tilde{\nu}_{max}$ = 1748 (CO) cm^{-1}. ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 5.30−5.25 (m, 1 H, 4-H), 5.21 (dd, $^3J_{2,1}$ = 7.1, $^3J_{2,3}$ = 9.6 Hz, 1 H, 2-H), 5.05 (dd, $^3J_{3,4}$ = 3.5, $^3J_{3,2}$ = 9.5 Hz, 1 H, 3-H), 4.46 (d, $^3J_{1,2}$ = 6.9 Hz, 1 H, 1-H), 4.05 (dd, $^3J_{5e,4}$ = 3.0, $^3J_{5e,5a}$ = 13.1 Hz, 1 H, 5-H_e), 3.92−3.73 (m, 4 H, 2 × CH_2), 3.66 (dd, $^3J_{5a,4}$ = 1.6, $^3J_{5a,5e}$ = 13.2 Hz, 1 H, 5-H_a), 2.15 (s, 3 H, CH_3CO), 2.08 (s, 3 H, CH_3CO), 2.03 (s, 3 H, CH_3CO) ppm. MS (EI): m/z (%) = 319.1030 [M − H] {$C_{13}H_{19}O_9$ requires 319.1029}, 73 (100).

2′-(Mesyloxy)ethyl 2,3,4-Tri-*O*-acetyl-α-L-arabinopyranoside (11): Triethylamine (0.20 mL, 1.4 mmol) was added to a stirred solution of sugar **10** (41 mg, 0.13 mmol) in tetrahydrofuran (4 mL) at 0 °C under nitrogen, followed by dropwise addition of mesyl chloride (0.10 mL, 1.3 mmol) and the mixture was stirred at 0 °C for 2 d. The solvent was removed and the residue taken up in cold chloroform, washed with ice-cold water and dried with anhydrous sodium sulfate. Removal of the solvent afforded mesylate **11** as a clear oil in quantitative yield. ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 5.30−5.25 (m, 1 H, 4-H), 5.19 (dd, $^3J_{2,1}$ = 6.8, $^3J_{2,3}$ = 9.4 Hz, 1 H, 2-H), 5.04 (dd, $^3J_{3,4}$ = 3.4, $^3J_{3,2}$ = 9.3 Hz, 1 H, 3-H), 4.48 (d, $^3J_{1,2}$ = 6.8 Hz, 1 H, 1-H), 4.37 (t, ^3J = 4.5 Hz, 4 H, 2 × 2′-H), 4.13−4.00 (m, 3 H, 5-H_e and 2 × 1′-H), 3.67−3.61 (m, 1 H, 5-H_a), 3.04 (s, 1 H, SO_2CH_3), 2.14 (s, 3 H, CH_3CO), 2.08 (s, 3 H, CH_3CO), 2.01 (s, 3 H, CH_3CO) ppm.

Naphth-2-yl 2′,3′,4′-Tri-*O*-acetyl-α-L-arabinopyranoside (12)

Mitsunobu Conditions: Triphenylphosphane (27.9 mg, 106 μmol) and diethyl azodicarboxylate (16.5 μL, 105 μmol) in tetrahydrofuran (0.15 mL) were added to a stirred solution of sugar **5** (23.8 mg, 86.2 μmol) and naphth-2-ol (10.1 mg, 70.1 μmol) in tetrahydrofuran (0.5 mL) at 0 °C, and the mixture was stirred at 0 °C for 2.75 h. Removal of the solvent afforded a cream-yellow oily solid which, upon purification by flash chromatography (dichloromethane), afforded glycoside **12** as a white solid (10.8 mg, 39%); m.p. > 300 °C. ^{1}H NMR (200 MHz, $CDCl_3$, 27 °C): δ = 7.77 (dd, ^3J = 1.5, 6.1 Hz, 2 H, aryl), 7.40 (m, 4 H, aryl), 7.20 (dd, ^3J = 2.4, 8.9 Hz, 1 H, aryl), 5.50 (dd, $^3J_{2',1'}$ = 6.2, $^3J_{2',3'}$ = 8.7 Hz, 1 H, 2′-H), 5.37 (m, 1 H, 4′-H), 5.25 (d, $^3J_{1',2'}$ = 6.1 Hz, 1 H, 1′-H_a), 5.20 (dd, $^3J_{3',4'}$ = 3.3, $^3J_{3',2'}$ = 8.7 Hz, 1 H, 3′-H), 4.17 (dd, $^3J_{5e',4'}$ = 4.1, $^3J_{5e',5a'}$ = 12.8 Hz, 1 H, 5′-H_e), 3.81 (dd, $^3J_{5a',4'}$ = 2.1, $^3J_{5a',5e'}$ = 12.8 Hz, 1 H, 5′-H_a), 2.16 (s, 3 H, CH_3CO), 2.11 (s, 3 H, CH_3CO), 2.10 (s, 3 H, CH_3CO) ppm. MS(EI): m/z (%) = 402.1315 [M^+] {$C_{21}H_{22}O_8$ requires 402.1330}, 259 (13.8) [M − $OC_{10}H_7$], 144 (100) [$C_{10}H_7OH$], 43 (98.4) [CH_3CO].

Königs−Knorr Conditions: Silver carbonate (24.4 mg, 88 μmol) was added to a solution of sugar **4** (36.9 mg, 109 μmol) and naphth-2-ol (10.7 mg, 74 μmol) in diethyl ether/dichloromethane (1:1, 1.6 mL) over 4-Å molecular sieves. The mixture was stirred in the dark at room temperature for 25.5 h, passed through a layer of silica to remove the silver salts and the solvent removed to give the crude product (19.5 mg). Purification by flash chromatography twice (0 → 2% methanol in dichloromethane; 0 → 0.5% methanol in dichloromethane), gave compound **12** (2.7 mg, 9%), which showed similar spectral properties to that isolated using the Mitsunobu conditions.

Glycosyl Fluoride Route: Boron trifluoride−diethyl ether (50 μL, 400 μmol) was added to a solution of fluoride **6** (24 mg, 87 μmol) and naphth-2-ol (6.8 mg, 47 μmol) in acetonitrile (2 mL), and the mixture stirred under nitrogen for 7 d. The solvent was removed and the residue taken up in $CHCl_3$, washed with saturated $NaHCO_3$ and water, and dried with anhydrous sodium sulfate. Removal of the solvent gave the crude product (12.5 mg), which was purified by flash chromatography (0 → 1% methanol in dichloromethane) to afford impure **12** (3.1 mg) which was identified on comparison with authentic material prepared as described above.

9-(α-L-Arabinopyranosyl)-2-methyllipticinium Acetate (13): Celiptium (**2**) (20.0 mg, 59.5 μmol) was dissolved in distilled methanol (5.0 mL) and chilled to 0 °C with stirring under nitrogen. A solution of sodium methoxide in methanol (0.30 mL, 2.1 M, 630 μmol) was added, followed by glycosyl bromide **4** (105 mg, 310 μmol), and the reaction mixture stirred at 0 °C for 10 min. The solution

was diluted with diethyl ether (3.0 mL) and the precipitate retained. Mass spectral analysis of the crude product was consistent with the presence of glycoside **13**. ^{1}H NMR (400 MHz, CD_3OD, 27 °C): δ = 9.63 (1-H), 8.15 (3-H), 8.05 (4-H), 7.60 (10-H), 7.28 (7-H), 7.05 (8-H), 4.1−3.6 (m, H-sugars) ppm. MS (ES^+): *m*/*z* (%) = 409 (10) [M^+], 278 (100).

2-(α-L-Arabinopyranosyl)-9-(α-L-arabinopyranosyl)ellipticinium Acetate (14): Compound **3**[15] (14.5 mg, 24 μmol) was dissolved in methanol (8 mL) and chilled to 0 °C with stirring under nitrogen. A solution of sodium methoxide in methanol (0.255 mL, 0.94 M, 240 μmol) was added, followed by glycosyl bromide **4** (41.4 mg, 122 μmol), and the reaction mixture stirred at 0 °C for 15 min, then allowed to warm to room temperature for a further 10 min. The solvent was removed and the residue washed with dichloromethane to removed unchanged sugar starting material, yielding the crude product. This contained predominantly unchanged starting material **3**, with only trace amounts of the desired product **14** which could not be isolated in sufficient amounts to allow full characterisation. ^{1}H NMR (400 MHz, CD_3OD, 27 °C): δ = 9.66 (1-H), 8.22 (3-H), 7.90 (4-H), 7.90 (10-H), 7.50 (7-H), 7.30 (8-H) ppm. MS (ES^+): *m*/*z* (%) = 527 (11) [M^+], 395 (100) [M − arab$^+$].

9-(2′-[α-L-Arabinopyranosyl]ethyloxy)-2-methylellipticinium Acetate (15): Celiptium (**2**) (20.0 mg, 59.5 μmol) was dissolved in methanol (3.5 mL) and chilled to 0 °C with stirring under nitrogen. A solution of sodium methoxide in methanol (0.65 mL, 0.94 M, 610 μmol) was added, followed by a solution of mesylate **11** (119 mg, 297 μmol) in methanol (3.0 mL), and the solution stirred under nitrogen at 0 °C for 20 min. The solvent was removed, the residue washed with dichloromethane to remove unchanged sugar starting materials and the resultant crude product was purified by reverse phase HPLC (40−85% methanol in 25 mM ammonium acetate over 45 min). Integration of the analytical HPLC trace (254 nm) showed the *O*-glycoside **15** was the major product present (ca. 25%). Purification of a portion of the material gave a number of mixed fractions in addition to pure **15** which was isolated as an orange solid (1.8 mg, 6%). UV (H_2O): λ_{max} (ε) = 210 (14300), 247 (17900), 278 (16800), 307 (27300), 379 (3700), 439 (2300) nm (dm^3 mol^{-1} cm^{-1}). ^{1}H NMR (400 MHz, CD_3OD, 27 °C): δ = 9.85 (s, 1 H, 1-H), 8.40 (d, $^3J_{3,4}$ = 7.2 Hz, 1 H, 3-H), 8.26 (dd, $^3J_{4,1}$ = 1.2, $^3J_{4,3}$ = 7.2 Hz, 1 H, 4-H), 7.96 (d, $^3J_{10,8}$ = 2.3 Hz, 1 H, 10-H), 7.56 (d, $^3J_{7,8}$ = 8.8 Hz, 1 H, 7-H), 7.34 (dd, $^3J_{8,10}$ = 2.4, $^3J_{8,7}$ = 8.8 Hz, 1 H, 8-H), 4.49 (s, 3 H, 2-CH_3), 4.39 (d, $^3J_{1'',2''}$ = 6.9 Hz, 1 H, 1′′-H), 4.36 (t, $^3J_{2',1'}$ = 4.8 Hz, 2 H, −OCH_2CH$_2$OAr), 4.23 and 4.01 (2 dt, $^3J_{1',2'}$ = 4.8, $^3J_{1A',1B'}$ = 11.4 Hz, (2 × 1 H, −OCH$_2$CH_2OAr), 3.93 (dd, $^3J_{5e'',4''}$ = 3.1, $^3J_{5e'',5a''}$ = 12.5 Hz, 1 H, 5′′-H$_e$), 3.86−3.84 (m, 1 H, 4′′-H), 3.67 (dd, $^3J_{2'',1''}$ = 6.8, $^3J_{2'',3''}$ = 8.9 Hz, 1 H, 2′′-H), 3.61 (dd, $^3J_{5a'',4''}$ = 1.6, $^3J_{5a'',4''}$ = 12.5 Hz, 1 H, 5′′-H$_a$), 3.58 (dd, $^3J_{3'',4''}$ = 3.5, $^3J_{3'',2''}$ = 8.9 Hz, 1 H, 3′′-H), 3.33 (s, 3 H, 11-CH_3), 2.87 (s, 3 H, 5-CH_3), 1.90 (s, 3 H, $CH_3CO_2^-$) ppm. MS (ES^+): *m*/*z* = 453.2019 [M^+] {$C_{25}H_{29}N_2O_6$ requires 453.2026}.

9-[2′-(α-L-Arabinopyranosyl)ethyloxy]-2-(α-L-arabinopyranosyl)-ellipticinium Acetate (16): Compound **3** (70 mg, 0.12 mmol) was dissolved in methanol (25 mL) and chilled to 0 °C with stirring under nitrogen. A solution of sodium methoxide in methanol (1.1 mL, 1.075 M, 1.2 mmol) was added and the mixture stirred for 10 min. A solution of mesylate **11** (231 mg, 0.58 mmol) in methanol (20 mL) was added, and the solution stirred at 0 °C for 30 min. The solvent was removed and the residue washed with dichloromethane to remove unchanged sugar starting material. The crude product was dissolved in water and purified by reverse phase HPLC (40−85% methanol in 25 mM ammonium acetate over 45 min followed by 40−70% methanol in 25 mM ammonium acetate over 60 min) to give a number of impure samples containing **16**. Pure compound **16** was isolated from several fractions as an orange solid (2.0 mg, 3%). UV (H_2O): λ_{max} (ε) = 211 (6900), 211 (6900), 249 (9600), 314 (16500), 454 (1700) nm (dm^{3}mol^{-1}cm^{-1}). ^{1}H NMR (400 MHz, CD_3OD, 27 °C): δ = 9.87 (s, 1 H, 1-H), 8.48 (dd, $^3J_{3,1}$ = 1.2, $^3J_{3,4}$ = 7.5 Hz, 1 H, 3-H), 8.30 (d, $^3J_{4,3}$ = 7.5 Hz, 1 H, 4-H), 7.74 (s, 1 H, 10-H), 7.44 (d, $^3J_{7,8}$ = 8.6 Hz, 1 H, 7-H), 7.22 (dd, 1 H, $^3J_{8,10}$ = 2.0, $^3J_{8,7}$ = 8.5 Hz, 8-H), 5.69 (d, $^3J_{1''',2'''}$ = 8.7 Hz, 1 H, 1′′′-H), 4.42 (d, $^3J_{1'',2''}$ = 7.0 Hz, 1 H, 1′′-H), 4.34−4.24 (m, 4 H, 2 × 2′-H and 1′-H$_A$ and 5′′′-H$_e$), 4.12−4.09 (m, 2 H), 4.05−4.01 (m, 1 H, 1′-H$_B$), 3.99−3.97 (m, 1 H), 3.92 (dd, $^3J_{5e'',4''}$ = 3.0, $^3J_{5e'',5a''}$ = 17.7 Hz, 1 H, 5′′-H$_e$), 3.88−3.86 (m, 2 H, 4′′-H and 1 H), 3.70 (dd, $^3J_{2'',1''}$ = 7.0, $^3J_{2'',3''}$ = 8.9 Hz, 1 H, 2′′-H), 3.64 (dd, $^3J_{5a'',4''}$ = 1.6, $^3J_{5a'',5e''}$ = 14.4 Hz, 1 H, 5′′-H$_a$), 3.61 (dd, $^3J_{3'',4''}$ = 3.4, $^3J_{3'',2''}$ = 8.9 Hz, 1 H, 3′′-H), 3.21 (s, 3 H, 11-CH_3), 2.77 (s, 3 H, 5-CH_3), 1.92 (s, 3 H, $CH_3CO_2^-$) ppm. MS (ES^+): *m*/*z* = 571.2271 [M^+] {$C_{29}H_{35}N_2O_{10}$ requires 571.2292}.

Ellipticin-9-yl 2′-(2′′,3′′,4′′-Tri-*O*-acetyl-α-L-arabinopyranosyl)glycolate (17): Acid **8** (17.5 mg, 52 μmol), *N*,*N*′-dicyclohexylcarbodiimide (13.1 mg, 63 μmol) and 1-hydroxybenzotriazole (7.4 mg, 55 μmol) were dissolved in DMF (1.5 mL) and stirred at room temperature for 4.5 h. 9-Hydroxyellipticine (**1**) (11.6 mg, 44 μmol) and (dimethylamino)pyridine (0.5 mg, 4 μmol) were added and stirring was continued for 42.5 h. The solvent was removed and the residue taken up in dichloromethane. The organic soluble portions were combined and the solvent removed to give the crude product, which was purified by flash chromatography (0 → 20% methanol in dichloromethane) to give compound **17** as a yellow solid (4.3 mg, 17%). ^{1}H NMR (400 MHz, CD_3OD, 27 °C): δ = 9.58 (s, 1 H, 1-H), 8.34 (d, $^3J_{3,4}$ = 6.2 Hz, 1 H, 3-H), 8.11 (d, 1 H, $^3J_{10,8}$ = 2.0 Hz, 10-H), 7.98 (d, $^3J_{4,3}$ = 6.1 Hz, 1 H, 4-H), 7.54 (d, $^3J_{7,8}$ = 8.7 Hz, 1 H, 7-H), 7.32 (dd, 1 H, $^3J_{8,10}$ = 2.1, $^3J_{8,7}$ = 8.7 Hz, 8-H), 5.32 (d, ^3J = 1.3 Hz, 1 H, 3′′-H), 5.24−5.17 (m, 2 H, 2′′-H and 4′′-H), 4.83 (d, $^3J_{1'',2''}$ = 6.6 Hz, 1 H, 1′′-H), 4.67 (ABq, ^3J = 4.8 Hz, 2 H, 2 × 2′-H,), 4.08 (dd, $^3J_{5e'',4''}$ = 3.0, $^3J_{5e'',5a''}$ = 13.2 Hz, 1 H, 5′′-H$_e$), 3.86 (dd, $^3J_{5a'',4''}$ = 1.4, $^3J_{5a'',5e''}$ = 13.3 Hz, 1 H, 5′′-H$_a$), 3.18 (s, 3 H, 11-CH_3), 2.76 (s, 3 H, 5-CH_3), 2.16 (s, 3 H, CH_3CO), 2.06 (s, 3 H, CH_3CO), 2.01 (s, 3 H, CH_3CO) ppm. MS (EI): *m*/*z* (%) = 578.1893 [M^+] {$C_{30}H_{30}N_2O_{10}$ requires 578.1901}, 262 (100) [**1**]. MS (ES^+): *m*/*z* (%) = 579 (87) [M + H^+], 239 (100).

DNA-Binding Studies: Water was purified using a Millipore Alpha-Q system. Stock solutions of calf thymus DNA (ct-DNA) (1500 μM) and separate solutions of the ellipticines **2**, **3**, **15** and **16** (800 μM) were prepared in water. The DNA concentration in bases was determined spectroscopically using ε_{258} = 6600 mol^{-1} dm^3 cm^{-1}. UV and LD titrations were performed by preparing a solution of DNA (500 μM) and adding aliquots of both the DNA stock and ellipticine stock solutions to give solutions with a constant DNA concentration and the following ratios of ellipticine/DNA: 1:50, 1:40, 1:30, 1:20, 1:10 and 1:5. Titration solutions (pH = 7) contained NaCl (20 mM) and Na_2HPO_4 (1 mM).

UV Absorbance: UV spectra were recorded with a Cary 4E UV/Vis/NIR spectrophotometer at 20 °C in a 1-cm quartz cuvette. Blank spectra of the relevant solvent were collected and subtracted from the sample spectra. UV spectra of the ellipticines (40 μM) in water were also acquired.

Linear Dichroism (LD): LD spectra were measured with a Jasco J-710 spectropolarimeter, adapted for LD measurements. Orientation of the ellipticine/DNA samples was achieved in a flow Couette cell[34] with an inner rotating cylinder with a base plate adapted for the smaller compartment. The experimental path length was 1 mm

and the rotation speed was ca. 1000 rpm for all samples. A background spectrum of phosphate buffer, at the same rotation speeds as the samples, was collected and subtracted from all spectra. Data were collected with a response time of 0.5 s, at a speed of 1000 nm min^{-1} and a step resolution of 0.5 nm. Data were averaged over 16 acquisitions for all spectra. LD^r spectra were calculated using $LD^r(\lambda) = LD(\lambda)/A(\lambda)$, after adjustment of absorbance values due to different path lengths.

Acknowledgments

Financial assistance from the Sydney University Cancer Research Fund (M. M. H.) and the award of a University of Sydney Postgraduate Scholarship (A. R. G.) are gratefully acknowledged.

[1] W. Goodwin, A. F. Smith, E. C. Horning, *J. Am. Chem. Soc.* **1959**, *81*, 1903−1908.
[2] L. K. Dalton, S. Demerac, B. C. Elmes, J. W. Loder, J. M. Swan, T. Teitei, *Aust. J. Chem.* **1967**, *20*, 2715−2727.
[3] J.-B. Le Pecq, N. Dat-Xuong, C. Gosse, C. Paoletti, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 5078−5082.
[4] C. Paoletti, J.-B. Le Pecq, N. Dat-Xuong, P. Juret, H. Garnier, J.-L. Amiel, J. Rouesse, *Recent Res. Cancer Res.* **1980**, *74*, 107−123.
[5] J. Rouëssé, M. Spielmann, F. Turpin, T. Le Chevalier, M. Azab, J. M. Mondésir, *Eur. J. Cancer* **1993**, *29A*, 856−859.
[6] M. Ohashi, T. Oki, *Exp. Opin. Ther. Patents* **1996**, *6*, 1285−1294.
[7] C. Auclair, *Arch. Biochem. Biophys.* **1987**, *259*, 1−14.
[8] P. Fossé, B. René, M. Charra, C. Paoletti, J.-M. Saucier, *Mol. Pharmacol.* **1992**, *42*, 590−595.
[9] E. Multon, J.-F. Riou, D. LeFevre, J.-C. Ahomadegbe, G. Riou, *Biochem. Pharmacol.* **1989**, *38*, 2077−2086.
[10] P. Formisyn, F. Pautet, C. Tran Minh, J. Bourgois, *J. Pharm. Biomed. Anal.* **1992**, *10*, 427−436.
[11] A. H. Elcock, A. Rodger, W. G. Richards, *Biopolymers* **1996**, *39*, 309−326.
[12] M. A. Ismail, K. J. Sanders, G. C. Fennell, H. C. Latham, P. Wormell, A. Rodger, *Biopolymers* **1998**, *46*, 127−143.
[13] M. M. Harding, A. R. Grummitt, *Mini-Rev. Med. Chem.*, in press.
[14] T. Honda, M. Inoue, M. Kato, K. Shima, T. Shimamoto, *Chem. Pharm. Bull.* **1987**, *35*, 3975−3978.
[15] T. Honda, M. Kato, M. Inoue, T. Shimamoto, K. Shima, T. Nakanishi, T. Yoshida, T. Noguchi, *J. Med. Chem.* **1988**, *31*, 1295−1305.
[16] T. Shimamoto, S. Imajo, T. Honda, S. Yoshimura, M. Ishiguro, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1331−1334.
[17] S. Walker, K. G. Valentine, D. Kahne, *J. Am. Chem. Soc.* **1990**, *112*, 6428−6429.
[18] D. Kahne, *Chem. Biol.* **1995**, *2*, 7−12.
[19] K. C. Nicolaou, K. Ajito, H. Komatsu, B. M. Smith, P. Bertinato, L. Gomez-Paloma, *Chem. Commun.* **1996**, 1495−1496.
[20] E. Palomino, *Adv. Drug Deliv. Rev.* **1994**, *13*, 311−323.
[21] L. F. Tietze, T. Feuerstein, A. Fecher, F. Haunert, O. Panknin, U. Borchers, I. Schuberth, F. Alves, *Angew. Chem. Int. Ed.* **2002**, *41*, 759−761.
[22] R. V. Stick, *Carboydrates − The Sweet Molecules of Life*, Academic Press, Salisbury, **2001**.
[23] P. Potier, *Chem. Soc. Rev.* **1992**, 113−119.
[24] V. P. Miller, D.-y. Yang, T. M. Weigel, O. Han, H.-w. Liu, *J. Org. Chem.* **1989**, *54*, 4175−4188.
[25] W. R. Roush, X.-F. Lin, *J. Am. Chem. Soc.* **1995**, *117*, 2236−2250.
[26] V. Magnus, D. Vikic-Topic, S. Iskric, S. Kveder, *Carbohydr. Res.* **1983**, *114*, 209−224.
[27] R. R. Schmidt, W. Kinzy, *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 21−123.
[28] P. J. Garegg, *Adv. Carbohydr. Chem. Biochem.* **1997**, *52*, 179−205.
[29] B. Dugue, B. Meunier, C. Paoletti, *Eur. J. Med. Chem. Chim. Ther.* **1983**, *18*, 551−554.
[30] K. Yoshida, N. Iino, I. Koga, *Chem. Pharm. Bull.* **1975**, *32*, 1759−1763.
[31] N. Harada, K. Ozaki, K. Oda, N. Nakanishi, M. Ohashi, T. Hashiyama, K. Tsujihara, *Chem. Pharm. Bull.* **1997**, *45*, 1156−1162.
[32] B. Nordén, T. Kurucsev, *J. Mol. Recognit.* **1994**, *7*, 141−156.
[33] B. Nordén, M. Kubista, T. Kurucsev, *Q. Rev. Biophys.* **1992**, *25*, 51−170.
[34] A. Rodger, *Methods Enzymol.* **1993**, *226*, 232−258.

Received July 19, 2002
[O02401]

Publication 83
Nucleic Acids Research, 2003, Vol. 31, No. 8 ***2097–2107***
DOI: 10.1093/nar/gkg316

Biophysical and biological properties of quadruplex oligodeoxyribonucleotides

Virna Đapić, Vedra Abdomerović, Rachel Marrington[1], Jemma Peberdy[1], Alison Rodger[1], John O. Trent and Paula J. Bates*

Department of Medicine, James Graham Brown Cancer Center, University of Louisville, Louisville, KY 40202, USA and [1]Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK

Received January 9, 2003; Revised February 12, 2003; Accepted February 21, 2003

ABSTRACT

Single-stranded guanosine-rich oligodeoxyribonucleotides (GROs) have a propensity to form quadruplex structures that are stabilized by G-quartets. In addition to intense speculation about the role of G-quartet formation *in vivo*, there is considerable interest in the therapeutic potential of quadruplex oligonucleotides as aptamers or non-antisense antiproliferative agents. We previously have described several GROs that inhibit proliferation and induce apoptosis in cancer cell lines. The activity of these GROs was related to their ability to bind to a specific cellular protein (GRO-binding protein, which has been tentatively identified as nucleolin). In this report, we describe the physical properties and biological activity of a group of 12 quadruplex oligonucleotides whose structures have been characterized previously. This group includes the thrombin-binding aptamer, an anti-HIV oligonucleotide, and several quadruplexes derived from telomere sequences. Thermal denaturation and circular dichroism (CD) spectropolarimetry were utilized to investigate the stability, reversibility and ion dependence of G-quartet formation. The ability of each oligonucleotide to inhibit the proliferation of cancer cells and to compete for binding to the GRO-binding protein was also examined. Our results confirm that G-quartet formation is essential for biological activity of GROs and show that, in some cases, quadruplex structures formed in the presence of potassium ions are significantly more active than those formed in the presence of sodium ions. However, not all quadruplex structures exhibit antiproliferative effects, and the most accurate factor in predicting biological activity was the ability to bind to the GRO-binding protein. Our data also indicate that the CD spectra of quadruplex oligonucleotides may be more complex than previously thought.

INTRODUCTION

It has long been recognized that G-rich nucleic acid sequences can adopt intermolecular or intramolecular quadruplex structures that are stabilized by the presence of G-quartets (Fig. 1A). There is currently substantial interest in the potential roles of quadruplex formation *in vivo* because many biologically important G-rich sequences are capable of forming G-quartets under physiological conditions *in vitro* (1–7). In addition, the number of reports describing specific G-quartet-binding proteins is now considerable [reviewed in Shafer and Smirnov (8)]. Telomerase, a promising molecular target for cancer therapy because it is expressed in almost all tumor cells but absent in most somatic tissues (9–11), has sparked additional interest in quadruplexes because stabilization of telomeres by G-quartet-interactive compounds may inhibit the activity of this enzyme (12–15).

G-quartet formation has also been implicated in the non-antisense antiproliferative effects of G-rich oligonucleotides (GROs). In several cases, the biological effects of oligonucleotides designed as antisense agents were found to be unrelated to inhibition of target protein expression, but instead were associated with the formation of G-quartet structures (16–19). Although disadvantageous for antisense activity, these quadruplex effects may have therapeutic utility. We recently have reported on 3′-protected G-rich phosphodiester oligodeoxyribonucleotides, specifically GRO29A, that have impressive antiproliferative activity when added to cancer cell lines (20–22). Active GROs were shown to be stabilized by G-quartet structures that are nuclease resistant, so that even a 3′-unmodified phosphodiester analog of GRO29A was undegraded in serum-containing cell culture medium for several days (22). The precise molecular mechanism of GRO action is not yet known, but appears to be related to the ability of oligonucleotides to bind to a specific protein, which was identified as nucleolin or a nucleolin-like protein (20). Binding of nucleolin to other G-quartet-forming sequences such as telomeric DNA, immunoglobulin switch regions and ribosomal genes has also been reported (2,5,23,24). Treatment of tumor cells with GROs was found to inhibit cell cycle progression by specifically interfering with DNA replication, whereas GRO-treated normal skin cells exhibited minimal perturbation of the cell cycle (21).

*To whom correspondence should be addressed at University of Louisville, 204B Baxter Biomedical Research Building, 570 South Preston Street, Louisville, KY 40202, USA. Tel: +1 502 852 2432; Fax: +1 502 852 2356; Email: paula.bates@louisville.edu
Correspondence may also be addressed to John O. Trent. Tel: +1 502 852 2194; Fax: +1 502 852 2195; Email: john.trent@louisville.edu

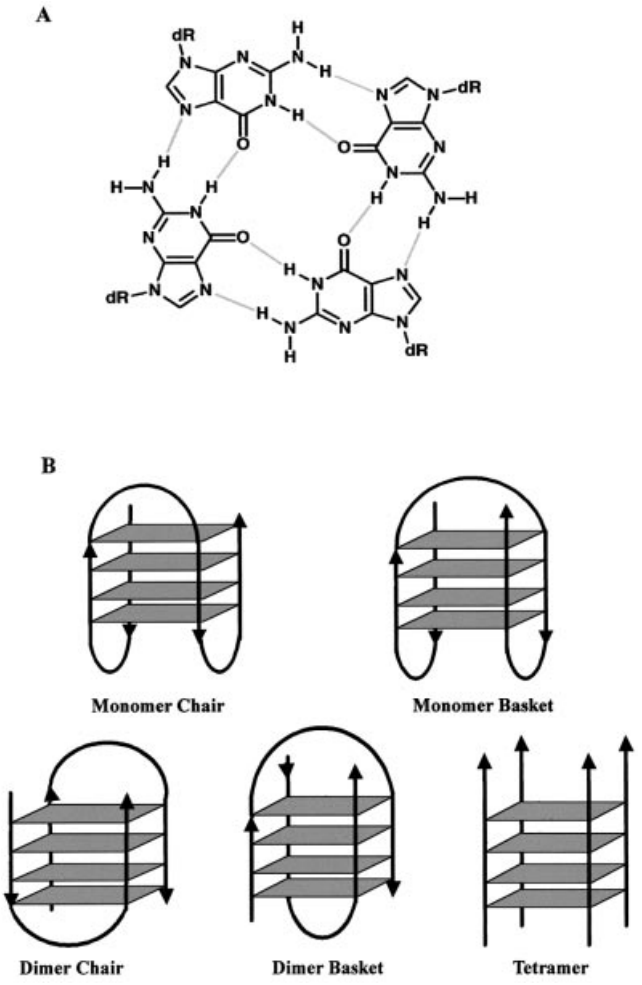


Figure 1. Schematic presentation of G-quartet structures. (**A**) G-quartet. (**B**) Molecularity and loop orientation of quadruplexes

GROs can form a variety of possible quadruplex structures, depending on both thermodynamic and kinetic considerations. The structures formed can be influenced by oligonucleotide base sequence and concentration, as well as the conditions (temperature and buffer) used for annealing, especially the presence of monovalent cations such as K^+ and Na^+. Quadruplexes can be formed by one, two or four molecules of oligonucleotide, which are referred to as monomer, dimer and tetramer structures, respectively (Fig. 1B). Monomer and dimer quadruplexes have been classified further based on the positioning of their loop regions into chair (lateral loop) or basket (diagonal loop) forms. The relative strand orientation (5′ to 3′ polarity) of the four strands of the quadruplex may be parallel (e.g. Fig. 1B, tetramer), antiparallel (Fig. 1B, monomer chair or dimer chair) or mixed (Fig. 1B, monomer basket or dimer basket).

The major aim of this study was to identify quadruplex features that are associated with antiproliferative activity in order to facilitate the design of GROs with optimal antitumor effects. A second goal was to investigate the physical properties of quadruplex oligonucleotides and their dependence on sodium and potassium cations. Therefore, we chose to study 12 quadruplex oligodeoxynucleotides whose structural characteristics had been determined previously by nuclear magnetic resonance (NMR) and/or X-ray crystallographic (XRC) techniques. The sequences and attributes of these oligonucleotides are shown in Table 1, and are described in more detail in the following section. The structures of many of these sequences have also been discussed in a recent excellent review (25). In many of the cases where both NMR and XRC structures are available, there is disagreement between the two techniques regarding the loop structure (basket or chair) or conformation (*syn* or *anti*) of the quartet guanosines (25), although the molecularities of the structures formed (tetramer, dimer or monomer) are in accord. These apparent discrepancies have been resolved to some degree by more recent structural evaluations (26,27), but it is possible that, in some cases, both structures may be correct and any disparities reflect the different experimental conditions used for quadruplex annealing.

Oligonucleotide KS-A, d(TGGGGT), is derived from the *Oxytricha nova* telomeric sequence, and previous work (both XRC and NMR) has shown that it forms a parallel tetrameric quadruplex in the presence of sodium ions (28–30). Oligonucleotide KS-B, d(GGTTGGTGTGGTTGG), is known as the thrombin-binding aptamer and was developed by an *in vitro* selection approach. It forms a monomeric chair structure that is strongly stabilized by potassium ions (31–33). The NMR and crystal structures are in accord concerning the chair conformation, but differ in the folding pattern (although it should be noted that the NMR analysis was carried out in the presence of potassium, whereas in the crystal structure the aptamer is in the presence of thrombin). However, re-evaluation of higher resolution XRC data (27) has also indicated that the diffraction data could fit equally well to the NMR solution structure. Oligonucleotide KS-C, d($G_4T_4G_4$), is derived from the *O.nova* telomere sequence and forms a dimeric hairpin quadruplex in the presence of K^+, Na^+ or NH_4^+ (34). This dimer adopts a chair conformation by XRC analysis (35), but exists in the basket conformation according to the NMR structure (36). A more recent XRC study indicates that a similar basket conformation can also be formed in the crystalline state (26). The KS-D sequence is a heptamer, d(GCATGCT), that forms a four-stranded structure that does not involve G-quartets. Rather, it forms a novel dimer structure stabilized by the non-classical base-pairing interactions of two folded molecules (37). Oligonucleotide KS-E, d(GCGGTTTGCGG), represents the fragile X gene repeat sequence, and has been shown to form a dimeric chair structure in the presence of sodium ions that contains two G-quartets and two G·C·G·C tetrads (3). Oligonucleotide KS-F, d(TAGG), is a short analog of the *Bombyx mori* telomere repeat and has been analyzed by a combined NMR–molecular dynamics approach. It forms an unusual structure containing two G-quartets, in which each strand has a parallel and an antiparallel neighbor (38). The KS-G sequence, d($G_3T_4G_3$), is another *O.nova* telomere sequence and forms a dimeric quadruplex structure in the presence of either K^+ or Na^+ (39,40), but is stabilized preferentially by potassium (41). NMR studies show that this quadruplex is in the basket form, but that the glycosidic conformations are different from the similar KS-C sequence. Both KS-H, d($G_4T_4G_4T_4G_4T_4G_4$), and KS-I, d($G_4T_2G_4T_2G_4T_2G_4$), which are derived from,

respectively, *O.nova* and *Tetrahymena thermophila* telomere sequences, form folded monomeric quadruplexes, but they have considerably different structures. The KS-H monomer is stabilized by four stacked G-quartets with two lateral T_4 loops and a diagonal central loop (42,43) so that each strand is adjacent to one parallel and one antiparallel strand (basket monomer). Due to the shorter sequence between the G_4 repeats, KS-I can only form three stacked G-quartets, and the folding of this structure is also quite different from that of KS-H. For KS-I, the quadruplex is formed with a lateral GTTG loop, a central lateral TTG loop and an unusual TT loop that spans from the top of the quadruplex to the bottom such that the fourth strand polarity is parallel to the adjacent first and third strands (44). T30695 is a potent anti-HIV oligonucleotide that has been analyzed by a combined NMR–molecular modeling approach and shown to form a potassium-stabilized chair monomer containing three stacked G-quartets (45). The sequences of GRO20A and GRO23A were based on a structure–activity study carried out by Marathias and Bolton (46). Using NMR analysis, these authors showed that the sequence similar to GRO20A (in the NMR study, some of the loop thymines were substituted by uracil) formed a well defined monomer basket quadruplex in the presence of sodium ions and did not change significantly upon addition of K^+. On the other hand, the spectrum of GRO23A was affected by the addition of potassium and was not consistent with a single well defined structure, so it was proposed that this oligonucleotide formed a complex mixture of chair and basket conformations.

MATERIALS AND METHODS

Oligodeoxyribonucleotides

All oligonucleotides had phosphodiester backbones and were synthesized with a 3′-propylamine modification. Oligonucleotides were purchased from Oligos Etc. (Wilsonville, OR) or were synthesized on a Beckman Oligo1000M synthesizer using reagents from Beckman, and 3′-propylamine CPG columns from Glen Research (Sterling, VA). We have shown previously that oligonucleotide activity is independent of the source of synthetic oligonucleotides. Oligonucleotides were resuspended in water, precipitated with butan-1-ol, washed with 70% ethanol, dried and resuspended in 10 mM Tris–HCl pH 7.4. They were then sterilized by filtration through a 0.2 μm filter, and the concentration was determined by UV spectroscopy. Extinction coefficients were calculated as the sum of individual nucleotide ε_{260} (11 700 $M^{-1}cm^{-1}$ for deoxyguanosine, 8800 $M^{-1}cm^{-1}$ for thymidine, 15 400 $M^{-1}cm^{-1}$ for deoxyadenosine, and 7300 $M^{-1}cm^{-1}$ for deoxycytidine). Oligonucleotides were diluted to 250 μM with sterile 10 mM Tris–HCl and stored in aliquots at –20°C. Each oligonucleotide was checked for integrity by 5′-radiolabeling followed by denaturing polyacrylamide gel electrophoresis. The 3′-propylamine was added to ensure that all oligonucleotides had similar stability in serum-containing medium, although we have shown that this modification does not greatly affect the activity, thermal stability (22) or circular dichroism (CD) spectrum (P.J.Bates and A.Rodger, unpublished observations) of active GROs.

Antiproliferative activity of G-rich oligonucleotides in buffers containing potassium and sodium

Oligonucleotides were diluted in 10 mM Tris–HCl pH 7.4 to give a final concentration of 150 μM. Samples were boiled for 5 min and placed on ice. Subsequently, sterile 2 M KCl or 2 M NaCl was added to produce a final 50 mM concentration of NaCl or KCl, and samples were incubated for 40 h at 60°C. These lengthy annealing conditions were used to ensure dissociation of any existing structures formed during oligonucleotide synthesis and resuspension, so that the final solutions contained the most thermodynamically stable forms in either potassium or sodium. This method therefore does not take into account the kinetic factors that may influence conformation in solution, and further studies will be required to establish which structural forms are prevalent under biological conditions.

HeLa cervical carcinoma cells were plated at a density of 10^3 cells/well in a 96-well plate in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, which had been heat inactivated for 30 min at 55°C. Annealed oligonucleotides (10 μl of 150 μM) were added to cells to give a final concentration of 10 μM oligonucleotide per well. Cells were placed in an incubator at 37°C in an atmosphere of 10% CO_2, and culture medium was not replaced for the duration of the experiment. Seven days after the addition of oligonucleotides, cell viability was determined using the MTT assay (47). Experiments were performed in triplicate, and bars represent the standard error of the data. Results were found to be reproducible, and treatment of cells with buffers containing amounts of NaCl or KCl equivalent to experimental samples was not significantly toxic (data not shown).

Thermal denaturation–renaturation using UV-visible spectroscopy

Oligonucleotides were resuspended in T_m buffer (140 mM KCl, 2.5 mM $MgCl_2$ and 20 mM Tris–HCl pH 8.0) at a final concentration ranging from 2 to 10 μM, depending on oligonucleotide length. Oligonucleotides were annealed in T_m buffer by boiling for 5 min, slow cooling to room temperature and overnight incubation at 4°C. Samples (1.6 ml, which filled the cuvette) were placed in a stoppered quartz cuvette of 1 cm path length and were allowed to reach ambient temperature before beginning each experiment. Thermal denaturation–renaturation was carried out using an Ultrospec 2000 UV/visible spectrophotometer equipped with a Peltier effect heated cuvette holder and temperature controller (Amersham Pharmacia Biotech). A temperature range of 25–95 or 20–90°C was used to monitor absorbance at 295 nm at a heating/cooling rate of 0.5°C/min. The melting and annealing temperatures were calculated as the temperature where the absorbance at 295 nm was halfway between the absorbance of the annealed species and the absorbance of the denatured species.

Circular dichroism spectroscopy

Oligonucleotides, at a final concentration of 5 μM, were resuspended in 10 mM sodium phosphate buffer pH 7.0 containing either 0.1 M NaCl or 0.1 M KCl. Samples were boiled for 5 min, placed on ice and annealed at 60°C for 56 h. CD spectra were collected on a Jasco J-715 spectropolarimeter

at 320–200 nm, using 16 scans at 100 nm/min, 1 s response time, 1 nm bandwidth. Cuvettes of 4 mm width with black quartz sides to mask the light beam were used for the measurements. A buffer baseline was collected in the same cuvette and subtracted from sample spectra. Final spectra were normalized to have zero ellipticity at 320 nm.

Protein-binding assay

An electrophoretic mobility shift assay (EMSA) to determine the ability of GROs to compete for binding to a specific GRO-binding protein (tentatively identified as nucleolin) was described previously (20). Briefly, a G-quartet-forming oligonucleotide representing the human telomere sequence, d(TTAGGGTTAGGGTTAGGGTTAGGG), referred to as 'TEL', was labeled with ^{32}P using T4 kinase. Labeled TEL (2×10^4 c.p.m. per reaction, ~1 nM final concentration) was pre-incubated alone or in the presence of an unlabeled competitor oligonucleotide (40 nM final concentration) for 30 min at 37°C. HeLa nuclear extracts (2.5 µg bandshift grade, Promega, Inc., Madison, WI) were added and samples were incubated for an additional 30 min at 37°C. Pre-incubation and binding reactions were carried out in buffer A [20 mM Tris–HCl pH 7.4, 140 mM KCl, 2.5 mM $MgCl_2$, 1 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride and 8% (v/v) glycerol]. Electrophoresis was carried out using 5% polyacrylamide gels in TBE buffer (90 mM Tris borate, 2 mM EDTA). Densitometry of autoradiograms was performed on the Personal Densitometer SI (Molecular Dynamics, Inc.) and image analyzed with ImageQuaNT™ software.

Non-denaturing polyacrylamide gel electrophoresis

Radiolabeled GRO29A (40 000 c.p.m.) was added to unlabeled oligonucleotide to give a final concentration of 300 µM. Samples were boiled for 5 min and placed on ice. An equal volume of buffer containing 10 mM Tris–HCl pH 7.4, supplemented with 100 mM KCl or 100 mM NaCl, or no salt was added to the sample. Therefore, samples contained final concentrations of 150 µM GRO29A, 10 mM Tris–HCl and 50 mM KCl or NaCl (or no additional salt), which is identical to samples used for the antiproliferative activity assay. Samples were annealed at 60°C for 48 h, and electrophoretic analysis was carried out using a 20% non-denaturing polyacrylamide gel in TBE buffer. Similar experiments were carried out with 1, 10 and 100 µM GRO29A (data not shown).

RESULTS

Cation-dependent antiproliferative activity of quadruplex oligonucleotides

To investigate how the antiproliferative effects of GROs depend on cations, and to determine whether other quadruplex oligonucleotides have similar activity, we annealed oligonucleotides in the presence of either potassium or sodium cations and tested their ability to inhibit the proliferation of HeLa cervical carcinoma cells.

The results of these experiments are shown in Figure 2, and the sequences and descriptions of all oligonucleotides used are outlined in Table 1. Oligonucleotides KS-A, KS-C, KS-D and GRO15B showed minimal antiproliferative activity in both NaCl and KCl buffers. Oligonucleotides KS-E, KS-F, KS-G and T30695 had intermediate activity that was not very different between NaCl and KCl buffers. The group of oligonucleotides that inhibited cellular proliferation by >50% was KS-B, KS-H, KS-I (in KCl), GRO20A, GRO23A (in KCl) and GRO29A. For some of these oligonucleotides (e.g. KS-I, GRO23A and GRO29A), there was a strong dependence on the annealing conditions, with K^+-annealed oligonucleotides having markedly enhanced activity compared with the Na^+-annealed sample.

With respect to developing a structure–activity relationship, these data do not appear to provide any definitive answers. Significant activity was observed for most, but not all, monomer quadruplexes, in both basket (e.g. GRO20A) and chair (e.g. KS-B) conformations. Perhaps most significant is the finding that not all quadruplex structures exhibit strong activity, suggesting that recognition of these molecules may be based on subtle features of the quadruplex, in accord with our previous studies (22). The present results also support our hypothesis (22) that the loop region is not a major determinant in quadruplex antiproliferative effects. In other words, although all active oligonucleotides in this study have loop structures, the number of loops, their base sequence and their conformation vary widely among the active molecules, making it unlikely that they contain a consensus domain for recognition.

UV thermal denaturation–renaturation

Our next aim was to determine whether biophysical signatures, such as the melting profile or CD spectrum, could predict the antiproliferative activity of quadruplex oligonucleotides. Thermal denaturation–renaturation studies can provide information about both the thermodynamic stability and kinetics of formation of secondary structures. We have used a spectroscopic method to detect G-quartet formation (20,22) that was first described by Mergny *et al.* (48). This technique relies on the observation that the absorbance of guanosines at 295 nm is higher when in a G-quartet structure than when denatured. In this study, we have determined the denaturation–renaturation profiles of quadruplex oligonucleotides and categorized them into three groups, as shown in Figure 3. The first category (NT) is exemplified by KS-A and showed no clear transition at 295 nm, indicating the absence of stable G-quartets (or possibly the presence of a structure with a melting temperature >90°C). The second category (R) was characterized by a clear and reversible transition where melting and annealing curves were almost superimposable, as in the case of KS-B. We interpreted this profile as reversible G-quartet formation with relatively fast kinetics. The third category (H) was typified by KS-C, and showed reversible melting with considerable hysteresis between melting and annealing curves, which is thought to reflect the slower kinetics (relative to the 'R' class) of quadruplex formation. The midpoints of the melting transition (T_{melt}) and the annealing transition (T_{anneal}) were calculated for all oligonucleotides and are shown in Table 2. Several trends can be observed by comparing Table 2 and Figure 2. First, with the exception of T30695 (which is discussed below), all oligonucleotides with the NT profile lack antiproliferative activity, supporting our previous observations (20) that G-quartet formation is necessary for activity. Secondly, not all oligonucleotides with profiles that indicate quadruplex

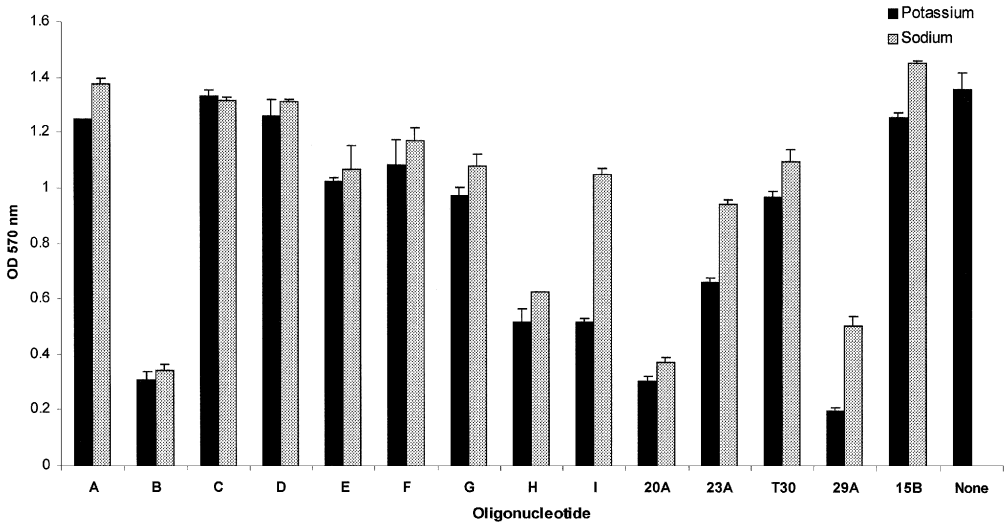


Figure 2. Cation dependence of antiproliferative activity. HeLa cervical carcinoma cells were either treated with a single dose of oligonucleotide (10 μM final concentration) annealed in buffer containing either 50 mM KCl or 50 mM NaCl, or received no oligonucleotide (None). Cell viability was assayed 7 days after addition of oligonucleotide using the MTT assay. All experiments were performed in triplicate, and bars represent the standard error of the data. In this figure, GRO20A, GRO23A, GRO29A, GRO15B and T30695 are abbreviated to 20A, 23A, 29A, 15B and T30 respectively. Dark bars represent oligonucleotides annealed in KCl and light bars represent oligonucleotides annealed in NaCl.

formation are active, confirming that G-quartet formation is not sufficient for activity (22). Finally, although all active oligonucleotides show thermal stability, there is no correlation between melting temperature and activity. For example, KS-B with a T_{melt} of 53°C has higher activity than GRO23A, which melts at ~85°C. Conversely, GRO29A with a T_{melt} of ~62°C has higher activity than GRO20A, which melts at 40°C.

It is not completely clear why some of the quadruplex-forming oligonucleotides in this study do not exhibit transitions in this assay. The utility of this approach in detecting the dissociation of parallel type quadruplexes such as KS-A has not been investigated extensively but, in the case of KS-D, KS-E and KS-F, it is most likely that these sequences do not form stable quadruplexes under the conditions used here. This hypothesis is supported by the fact that these oligonucleotides exhibit small ellipticities in their CD spectra (Fig. 4), which suggests the absence of any strong base stacking interactions. For oligonucleotide T30695, which has been shown previously to form an extremely stable quadruplex in the presence of potassium ions (45,49), it is more likely that the lack of a clear transition in this assay can be attributed a very high melting temperature. This is also supported by the CD of T30695, which shows a well defined spectrum with a large positive peak, indicative of strong base stacking interactions. In most respects, our results are consistent with other reports (50). Previous evaluations of four-stranded structures have shown that, generally, the kinetics of tetramer quadruplex formation are extremely slow, while monomeric quadruplexes form quickly and reversibly. The association rate for dimeric structures is also relatively fast, but melting profiles may show slight hysteresis (50).

Table 1. Sequences and structural properties of synthetic oligonucleotides

Oligonucleotide	Sequence	Description	References
KS-A (6mer)	d(TGGGGT)	Tetramer, telomere	(28–30,59)
KS-B (15mer)	d(GGTTGGTGTGGTTGG)	Monomer, thrombin aptamer	(27,31–33)
KS-C (12mer)	d(GGGGTTTTGGGG)	Dimer, telomere	(26,34–36,43)
KS-D (7mer)	d(GCATGCT)	Dimer, non-GQ	(37,60)
KS-E (11mer)	d(GCGGTTTGCGG)	Dimer, fragile X	(3)
KS-F (4mer)	d(TAGG)	Tetramer, telomere	(38)
KS-G (10mer)	d(GGGTTTTGGG)	Dimer, telomere	(39–41,61)
KS-H (28mer)	d(GGGGTTTTGGGGTTTTGGGGTTTTGGGG)	Monomer, telomere	(42,43,62)
KS-I (24mer)	d(TTGGGGTTGGGGTTGGGGTTGGGG)	Monomer, telomere	(44)
T30695 (16mer)	d(GGGTGGGTGGGTGGGT)	Monomer, anti-HIV	(45)
GRO20A (20mer)	d(GGTTTTGGTTTTGGTTTTGG)	Monomer, basket, experimental	(46)
GRO23A (23mer)	d(GGGGTTGGGGTGTGGGGTTGGGG)	Mixed basket/chair, experimental	(46)
GRO29A (29mer)	d(TTTGGTGGTGGTGGTTGTGGTGGTGGTGG)	Active GRO (proposed dimer)	(20–22)
GRO15B (15mer)	d(TTGGGGGGGGTGGGT)	Control GRO	(20–22)

Underlined bases form loop regions.

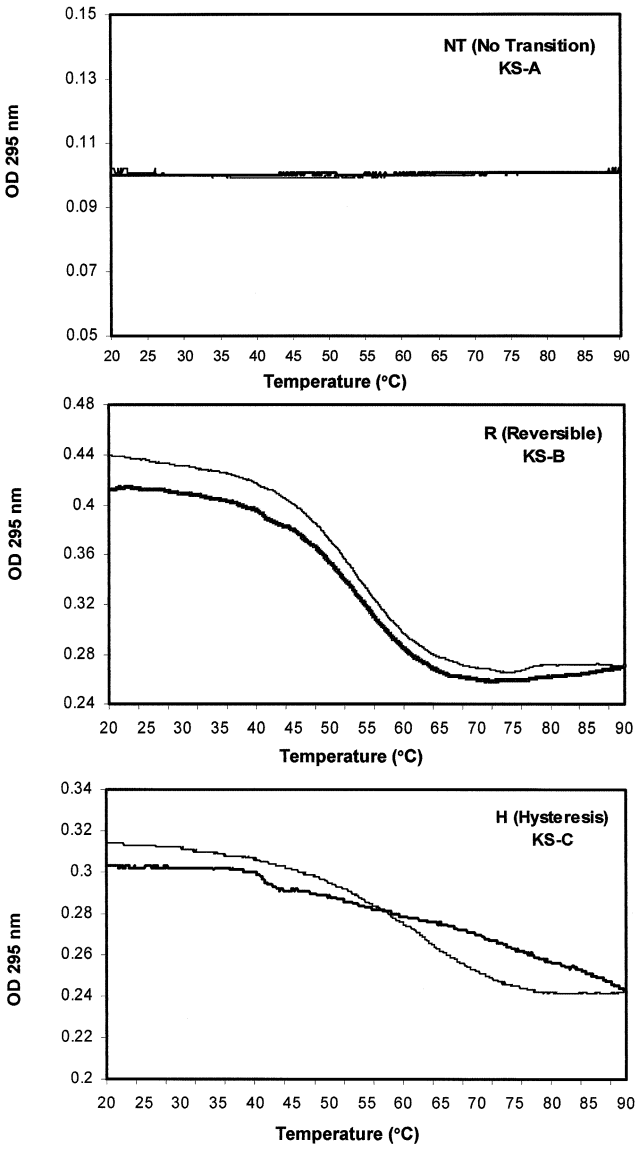


Figure 3. UV thermal denaturation–renaturation studies. Profiles were classified as NT, R or H, as shown. Oligonucleotides (2–10 μM final concentration) were annealed in T_m buffer (20 mM Tris–HCl pH 8.0, 140 mM KCl, and 2.5 mM $MgCl_2$) and absorbance was measured at 295 nm with a 1 cm path length. The melting curve is the lighter line; the annealing curve is the darker line. Melting and annealing temperatures are reported in Table 2.

CD spectropolarimetry

CD, which measures the differences between the absorbance of right-handed and left-handed circularly polarized light, has been used extensively to investigate the structure of helical molecules such as nucleic acids (51). The majority of publications have reported that 'folded' quadruplexes (monomers and dimers) have a CD spectrum characterized by a positive ellipticity maximum at 295 nm and a negative minimum at 265 nm, while the 'parallel' type (tetramers) have a positive maximum at 264 nm and a negative minimum at 240 nm (52–54). This distinction has become accepted by many researchers, and we previously were surprised to discover that the antiproliferative oligonucleotide GRO29A, which is thought to form a folded dimer quadruplex (22), exhibited a spectrum with a positive peak at 264 nm, typical of a parallel quadruplex. Therefore, the purpose of this part of the study was to clarify the relationship between CD spectral characteristics and quadruplex properties, and also to determine if any features in the CD spectra of quadruplexes were associated with biological activity. Figure 4 shows the CD spectra of all oligonucleotides annealed in buffers containing either 0.1 M KCl (thick line) or 0.1 M NaCl (thin line). As in the UV melting analysis, certain trends were observed, but there was no signature CD spectrum that was unequivocally linked with antiproliferative activity. Generally, oligonucleotides with weak activity (KS-A, KS-D, KS-E and KS-F) had small ellipticities with little difference between Na^+ and K^+ spectra. This suggested that only a small amount of quadruplex formation occurred under these conditions, consistent with the UV data. Those oligonucleotides whose CD signature was most different in form between NaCl and KCl buffers (KS-I and GRO23A) also exhibited clearly different antiproliferative activity in KCl and NaCl buffers. Most of the active oligonucleotides had the classical 'folded' CD spectrum (positive peak at 295 nm, negative at 265 nm). Some sequences exhibited this spectrum in both KCl and NaCl buffers, whereas for others the negative peak at 265 nm was only observed in sodium. However, it seems unlikely that the ability to exhibit this specific spectrum is important for biological activity because this spectral characteristic is associated with sodium ions, while activity is promoted by potassium ions. In addition, the spectrum corresponding to the most active oligonucleotide (GRO29A annealed in KCl, Fig. 6) is lacking this feature. In summary, the spectra suggest that there is not one structural reason for activity, but that activity is suggested by the presence of a number of spectral features. These include a large positive CD signal at either 264 or 295 nm, significant ellipticity in the 290–310 nm region (in contrast to duplex DNA, which lacks significant ellipticity at wavelengths >300 nm) and a large positive peak at 210 nm.

What is most striking about these data is that the CD spectra are often highly dependent on ionic conditions and cannot be categorized as simply 'parallel' or 'folded'. Many spectra (e.g. KS-I and GRO23A) contain positive peaks at both 264 and 295 nm. The appearance of these two peaks most probably indicates that more than one molecular species (probably both parallel and folded forms) is present in solution. The formation of multiple quadruplex species is another factor that adds to the difficulty in clearly defining quadruplex molecular structure, and may well contribute to the lack of simple correlation between structure and activity. The situation is made even more complex by the possibility that the predominant species in solution may be determined by kinetic considerations, as well as thermodynamic factors. Furthermore, an oligonucleotide (T30695) that has been well characterized as forming a monomeric folded quadruplex has a CD spectrum with a strong positive peak at 264 nm. In summary, while CD studies are clearly useful in establishing the presence of quadruplex structures, a much better understanding of the various contributions to quadruplex CD spectra will be required before CD data alone can be used to determine quadruplex molecular structure definitively. Until such results

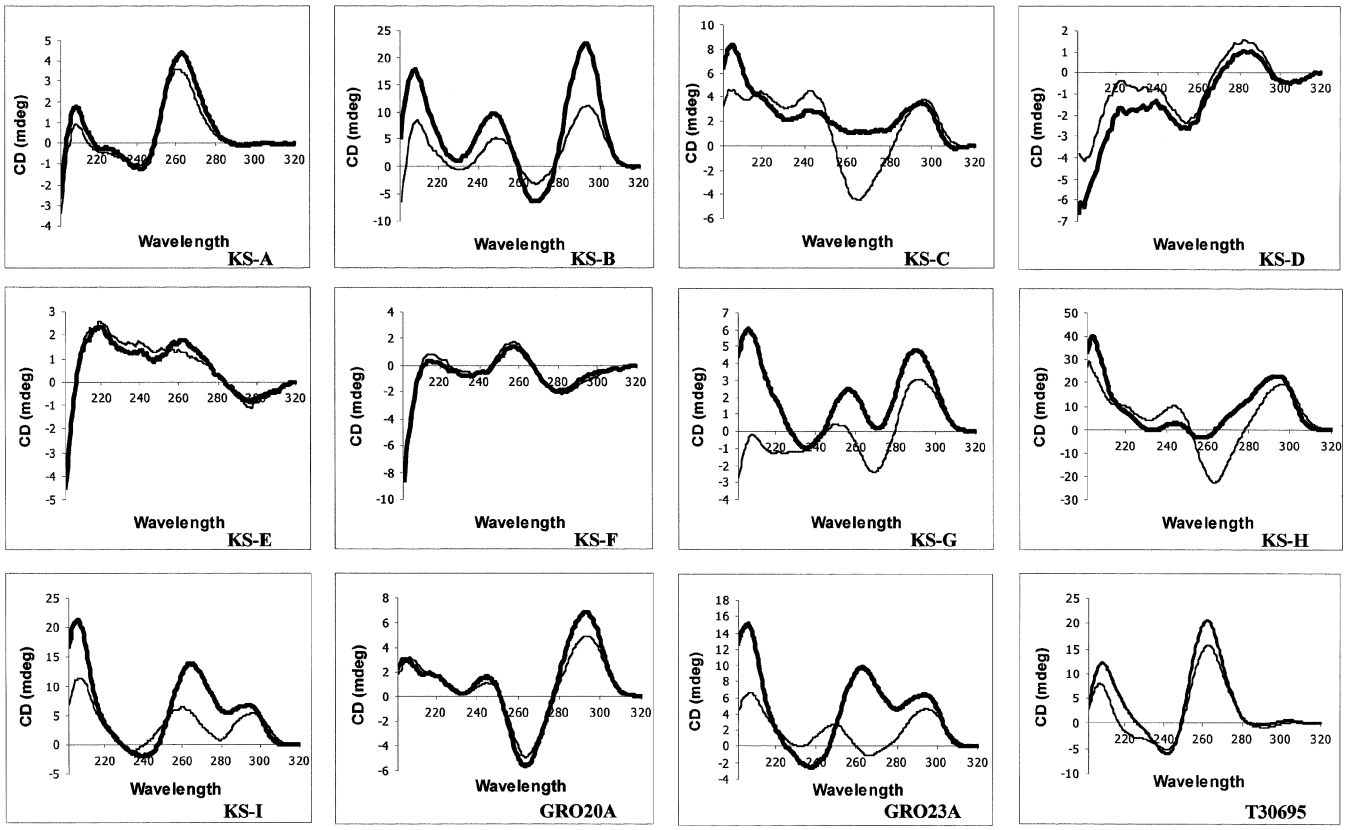


Figure 4. CD spectroscopy studies. CD spectra of oligonucleotides (5 μM final concentration) were obtained in the presence of either 0.1 M KCl (dark lines) or 0.1 M NaCl (light lines) at 25°C with a path length of 4 mm.

are available, researchers in this area will continue to rely on other biophysical methods to complement CD studies.

Binding of quadruplex oligonucleotides to GRO-binding protein

In previous work, we showed that the antiproliferative activity of GROs was related to their ability to compete with a ^{32}P-labeled telomere sequence oligonucleotide for binding to a specific cellular protein in a mobility shift assay (20,22). This specific GRO-binding protein was also recognized by a nucleolin antibody, and therefore we proposed that it was nucleolin or a similar protein (20). Figure 5A shows the relative ability of each quadruplex oligonucleotide to bind to the GRO-binding protein in a physiological buffer containing 140 mM KCl and 2.5 mM $MgCl_2$. In this assay, the ability of an oligonucleotide (which is unlabeled) to compete with labeled TEL for binding to GRO-binding protein (derived from HeLa extracts) is determined by the disappearance of the band indicated by an asterisk. There appears to be a good correlation between antiproliferative activity in KCl buffer and the ability of oligonucleotides to compete for the GRO-binding protein, with the most active oligonucleotides, namely GRO29A, KS-B, KS-H, KS-I and 20A (see Fig. 2), being most effective at competing for binding to GRO-binding protein. Figure 5B shows a plot of the relative intensity of the protein band in Figure 5A (note that in this assay, a high intensity indicates poor binding) versus the relative number of viable cells remaining after treatment with oligonucleotide annealed in KCl buffer (a high number indicates poor antiproliferative activity). The squared correlation coefficient (R^2) for the relationship between these variables is 0.82.

Implications for the structure of GRO29A

Another reason for carrying out biophysical studies of known quadruplex-forming oligonucleotides was to facilitate interpretation of the biophysical data regarding GRO29A and other active GROs. Previous molecular modeling studies (22) had predicted that GRO29A formed a dimeric quadruplex with lateral loops at opposite ends (chair form). The UV melting profile and CD spectra of GRO29A are shown in Figure 6. This oligonucleotide exhibits a fully reversible transition in T_m buffer (containing 140 mM KCl, 2.5 mM $MgCl_2$) with only slight hysteresis. This is consistent with the relatively fast kinetics of formation seen for both monomer and dimer quadruplexes (Table 2). The CD spectrum of GRO29A in KCl buffer shows a large positive peak at 264 nm with a small shoulder at 295 nm, and in NaCl solution the magnitude of the 264 nm peak is greatly reduced. Although the KCl spectrum (which, as Fig. 2 shows, represents the most active species) may have been interpreted previously as formation of a parallel tetramer quadruplex, our studies have indicated that the interpretation of CD data may not be as unambiguous as we had once thought. Thus, the peak at 264 nm does not necessarily exclude the possibility of a folded antiparallel

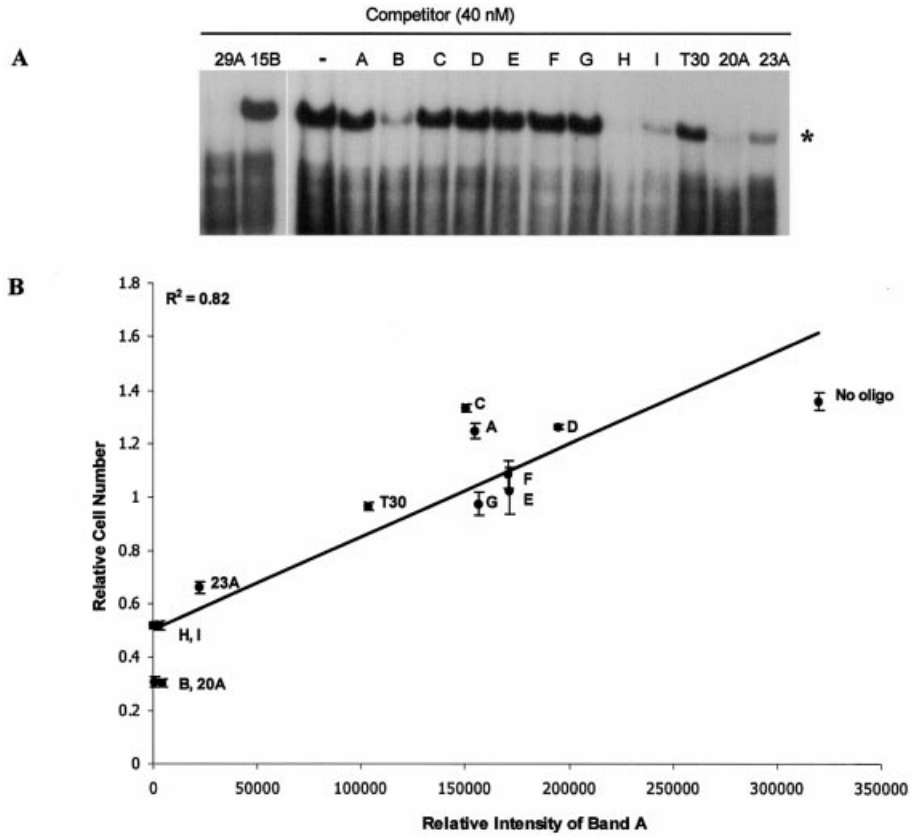


Figure 5. Protein-binding assay. (**A**) Competitive mobility shift assay showing the ability of unlabeled oligonucleotides to compete with radiolabeled TEL for binding to nuclear proteins. The band marked by an asterisk indicates the specific GRO-binding protein (thought to be nucleolin). The names of the oligonucleotides are abbreviated without the KS or GRO prefixes. (**B**) Plot of the relative intensity of the protein band in (A) versus the relative number of viable cells remaining after treatment with oligonucleotide annealed in KCl buffer. The squared correlation coefficient (R^2) for the relationship between these variables is 0.82. The names of the oligonucleotides are abbreviated without the KS or GRO prefixes.

quadruplex. A further possibility that should be considered is that GRO29A forms a structure that is folded, yet with strands in a parallel orientation. Such unusual structures have been reported recently for the human telomere sequence (55) and d(GGAGGAGGAGGAGGA) (56), and the latter structure does indeed exhibit a CD spectrum with a large positive peak at 264 nm (56).

In an attempt to clarify the molecularity of GRO29A, non-denaturing polyacrylamide electrophoresis was also performed after annealing GRO29A in various buffers (Fig. 6C). In the absence of monovalent cations, GRO29A migrates as two distinct bands on a native gel in TBE buffer, clearly indicating that multiple molecular species are present. Pre-annealing the oligonucleotide in 50 mM NaCl promotes the formation of the slower mobility species, whereas annealing in 50 mM KCl promotes the formation of the faster migrating species. This faster migrating species (which is most probably the active species) migrates parallel with denatured GRO29A and an unstructured oligonucleotide of the same length (data not shown). However, in the KCl sample, this lower band is unlikely to represent unstructured single strand because the melting temperature (in 50 mM KCl) was 49°C, whereas no clear melting transitions were seen in water or 50 mM NaCl. The appearance of bands in KCl buffer was independent of GRO29A concentration, whereas as in NaCl buffer, the intensity of the slower migrating band increased with increasing GRO29A concentration (for 1, 10 and 100 μM, data not shown). Although these electrophoretic data are far from equivocal, the mobility of the lower band suggests that the active species is likely to be a monomer (or, possibly, a very compact dimer). In summary, the biophysical data are not inconsistent with our previous molecular modeling prediction that GRO29A is a dimer (22), but could also be consistent with a monomeric structure. Further experiments are in progress to determine unequivocally the molecularity and structure of the most active form of GRO29A.

DISCUSSION

With growing evidence that G-quartet formation occurs *in vivo* and that such structures may be important therapeutic targets, there is an urgent need for a better understanding of quadruplex structures. NMR and crystallography are traditionally considered the 'gold standard' of structural techniques, but may be limited in their application to quadruplexes by experimental requirements for particular ionic conditions or strand concentrations, and the necessity to prepare a single species for structure determination. Molecular dynamics has

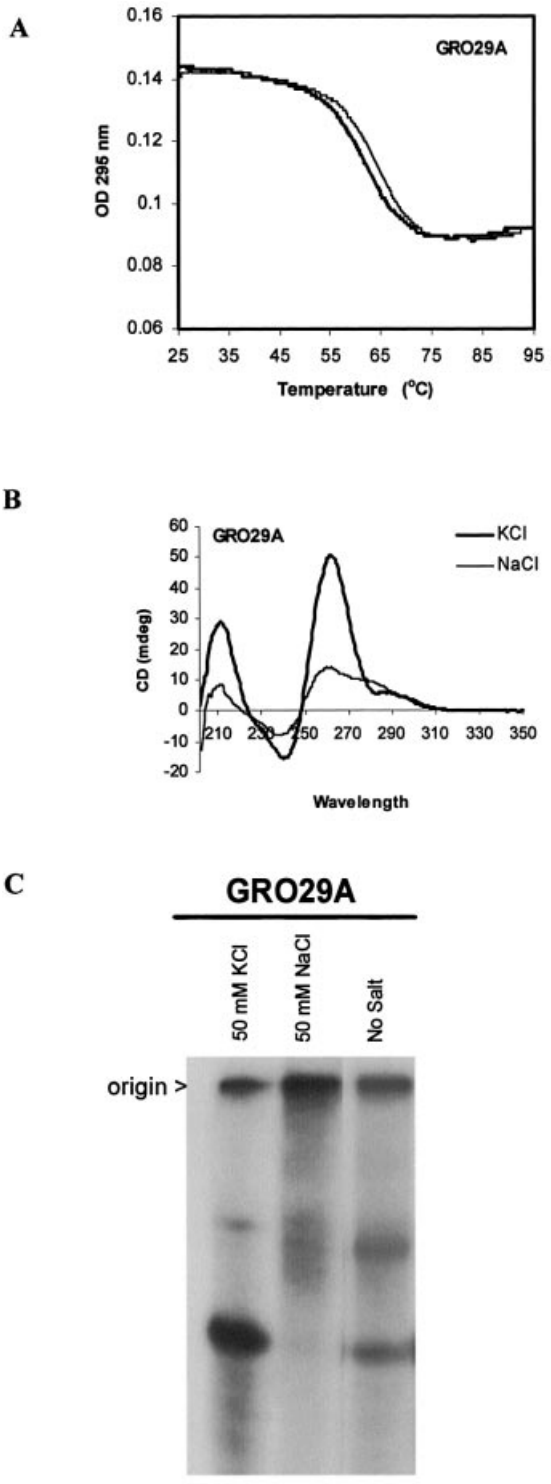


Figure 6. Structural characteristics of GRO29A. (**A**) UV thermal denaturation–renaturation profile (295 nm) of GRO29A (2 μM concentration, 1 cm path length). (**B**) CD spectra of GRO29A (5 μM concentration, 4 mm path length) in buffers containing KCl (dark line) or NaCl (light line). (**C**) Non-denaturing polyacrylamide electrophoresis of GRO29A (150 μM) annealed in 50 mM KCl, 50 mM NaCl or no salt.

also proved to be a powerful tool in examining quadruplex structures, but usually relies on biophysical data to generate starting structures or to confirm predictions. Biophysical methods that can be applied under a variety of experimental conditions to distinguish different G-quartet-containing structures would facilitate studies of quadruplex formation. Previously, biophysical methods such as UV melting, CD, calorimetry and NMR have been applied. These studies have yielded an abundance of thermodynamic and kinetic data that have been summarized comprehensively in a recent review (50). In spite of this, there is still no clear consensus regarding many features of G-quartet formation.

The aim of the present study was not to provide an exhaustive or quantitative analysis of biophysical properties, but to investigate simple relationships between the structure of quadruplexes, their biophysical properties and their antiproliferative activity. The major conclusion is that such simple relationships do not exist; rather the biophysical signatures (especially CD spectra) are exquisitely sensitive to small changes in quadruplex structure. In retrospect, this finding is unsurprising given that the form of the CD spectra will depend on the relative orientation and stacking of bases. These are not simply dependent on sequence, strand polarity or molecularity of the quadruplex, but affected by more complex factors such as glycosidic bond conformation and sugar pucker. The complexity of the spectra is enhanced further by the possibility that more than one quadruplex species may be present. The observation that a monomeric quadruplex (in this case, the anti-HIV oligonucleotide T30695) displays a classical 'parallel' CD spectrum with a positive peak at 264 nm is not novel (49,50,57), but should be emphasized because oversimplification of structure–spectrum relationships may lead to misinterpretation of quadruplex CD spectra. Although we cannot absolutely exclude the possibility that T30695 forms a parallel quadruplex under our conditions, this seems unlikely in light of previous studies of this sequence (45,49,57,58).

In our previous work, we have shown that G-quartet-forming oligonucleotides such as GRO29A have inhibitory effects on the proliferation of cancer cells (20–22). A previous study of backbone-modified GRO29A analogs (22) suggested that while G-quartet formation is necessary for activity, it is not sufficient, and binding to the GRO-binding protein is the primary determinant of activity. The ultimate goal of the present study was to extend our investigations to quadruplexes with characterized structures in order to identify features, such as molecularity and loop orientation, which are associated with non-antisense antiproliferative activity. We observed biological activity for some, but not all, of the previously characterized quadruplexes. Active structures included quadruplexes with a variety of conformations, suggesting that biological activity depends on subtle features of the quadruplex, rather than simply recognition of the G-quartet motif or a four-stranded structure. Biological activity requires both a reversible UV melting profile at 295 nm and the ability to compete for the GRO-binding protein (nucleolin) in the EMSA. The results presented in this report show that there is no simple recognition element associated with activity. This supports our hypothesis that protein binding (and thus biological activity) is mediated by recognition of the specific shape of the quadruplex groove, which will depend on many factors and could theoretically be similar in quadruplexes with very different sequences.

Table 2. Melting and annealing temperatures of oligonucleotides

Oligonucleotide	$T_{melting}$	$T_{annealing}$	Profile[a]	Molecularity	Activity[b]
KS-A	–	–	NT[a]	Tetramer	–
KS-B	52.5	53.0	R	Monomer	+
KS-C	68.5	62.8	H	Dimer	–
KS-D	–	–	NT	Dimer	–
KS-E	–	–	NT	Dimer	–
KS-F	–	–	NT	Tetramer	–
KS-G	45.6	41.4	R	Dimer	–
KS-H	65.3	66.8	R	Monomer	+
KS-I	86.1	82.8	H	Monomer	+
GRO20A	39.7	40.3	R	Monomer	+
GRO23A	85.6	83.8	R	Monomer	+
T30695	–	–	NT	Monomer	–
GRO29A	61.0	63.5	R	Proposed dimer	+
GRO15B	–	–	NT	No structure detected	–

[a]NT = no transition; R = reversible; H = hysteresis.
[b]Defined as >50% inhibition of HeLa cell proliferation at 10 μM in buffer containing 50 mM KCl.

ACKNOWLEDGEMENTS

The authors thank Dr Andrew N. Lane for helpful discussions during the preparation of this manuscript. This work was supported by the Department of Defense, the National Cancer Institute, the Commonwealth of Kentucky Research Challenge Trust and the Engineering and Physical Sciences Research Council (UK).

REFERENCES

1. Williamson,J.R. (1994) G-quartet structures in telomeric DNA. *Annu. Rev. Biophys Biomol. Struct.*, **23**, 703–730.
2. Dempsey,L.A., Sun,H., Hanakahi,L.A. and Maizels,N. (1999) G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D. A role for G–G pairing in immunoglobulin switch recombination. *J. Biol. Chem.*, **274**, 1066–1071.
3. Kettani,A., Kumar,R.A. and Patel,D.J. (1995) Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J. Mol. Biol.*, **254**, 638–656.
4. Simonsson,T., Pecinka,P. and Kubista,M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.*, **26**, 1167–1172.
5. Hanakahi,L.A., Sun,H. and Maizels,N. (1999) High affinity interactions of nucleolin with G–G-paired rDNA. *J. Biol. Chem.*, **274**, 15908–15912.
6. Catasti,P., Chen,X., Moyzis,R.K., Bradbury,E.M. and Gupta,G. (1996) Structure–function correlations of the insulin-linked polymorphic region. *J. Mol. Biol.*, **264**, 534–545.
7. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
8. Shafer,R.H. and Smirnov,I. (2001) Biological aspects of DNA/RNA quadruplexes. *Biopolymers*, **56**, 209–227.
9. Newbold,R.F. (1999) Telomerase as an anti-cancer drug target: will it fulfil its early promise? *Anticancer Drug Des.*, **14**, 349–354.
10. Lavelle,F., Riou,J.F., Laoui,A. and Mailliet,P. (2000) Telomerase: a therapeutic target for the third millennium? *Crit. Rev. Oncol. Hematol.*, **34**, 111–126.
11. Neidle,S. and Kelland,L.R. (1999) Telomerase as an anti-cancer target: current status and future prospects. *Anticancer Drug Des.*, **14**, 341–347.
12. Perry,P.J. and Jenkins,T.C. (1999) Recent advances in the development of telomerase inhibitors for the treatment of cancer. *Expert Opin. Invest. Drugs*, **8**, 1981–2008.
13. White,L.K., Wright,W.E. and Shay,J.W. (2001) Telomerase inhibitors. *Trends Biotechnol.*, **19**, 114–120.
14. Mergny,J.L., Mailliet,P., Lavelle,F., Riou,J.F., Laoui,A. and Helene,C. (1999) The development of telomerase inhibitors: the G-quartet approach. *Anticancer Drug Des.*, **14**, 327–339.
15. Bearss,D.J., Hurley,L.H. and Von Hoff,D.D. (2000) Telomere maintenance mechanisms as a target for drug development. *Oncogene*, **19**, 6632–6641.
16. Burgess,T.L., Fisher,E.F., Ross,S.L., Bready,J.V., Qian,Y.X., Bayewitch,L.A., Cohen,A.M., Herrera,C.J., Hu,S.S., Kramer,T.B. *et al.* (1995) The antiproliferative activity of c-myb and c-myc antisense oligonucleotides in smooth muscle cells is caused by a nonantisense mechanism. *Proc. Natl Acad. Sci. USA*, **92**, 4051–4055.
17. Anselmet,A., Mayat,E., Wietek,S., Layer,P.G., Payrastre,B. and Massoulie,J. (2002) Non-antisense cellular responses to oligonucleotides. *FEBS Lett.*, **510**, 175–180.
18. Benimetskaya,L., Berton,M., Kolbanovsky,A., Benimetsky,S. and Stein,C.A. (1997) Formation of a G-tetrad and higher order structures correlates with biological activity of the RelA (NF-κB p65) 'antisense' oligodeoxynucleotide. *Nucleic Acids Res.*, **25**, 2648–2656.
19. Saijo,Y., Uchiyama,B., Abe,T., Satoh,K. and Nukiwa,T. (1997) Contiguous four-guanosine sequence in c-myc antisense phosphorothioate oligonucleotides inhibits cell growth on human lung cancer cells: possible involvement of cell adhesion inhibition. *Jpn J. Cancer Res.*, **88**, 26–33.
20. Bates,P.J., Kahlon,J.B., Thomas,S.D., Trent,J.O. and Miller,D.M. (1999) Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *J. Biol. Chem.*, **274**, 26369–26377.
21. Xu,X., Hamhouyia,F., Thomas,S.D., Burke,T.J., Girvan,A.C., McGregor,W.G., Trent,J.O., Miller,D.M. and Bates,P.J. (2001) Inhibition of DNA replication and induction of S phase cell cycle arrest by G-rich oligonucleotides. *J. Biol. Chem.*, **276**, 43221–43230.
22. Dapic,V., Bates,P.J., Trent,J.O., Rodger,A., Thomas,S.D. and Miller,D.M. (2002) Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications. *Biochemistry*, **41**, 3676–3685.
23. Ishikawa,F., Matunis,M.J., Dreyfuss,G. and Cech,T.R. (1993) Nuclear proteins that bind the pre-mRNA 3′ splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. *Mol. Cell. Biol.*, **13**, 4301–4310.
24. Dickinson,L.A. and Kohwi-Shigematsu,T. (1995) Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential. *Mol. Cell. Biol.*, **15**, 456–465.
25. Keniry,M.A. (2001) Quadruplex structures in nucleic acids. *Biopolymers*, **56**, 123–146.
26. Haider,S., Parkinson,G.N. and Neidle,S. (2002) Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J. Mol. Biol.*, **320**, 189–200.
27. Padmanabhan,K. and Tulinsky,A. (1996) An ambiguous structure of a DNA 15-mer thrombin complex. *Acta Crystallogr. D*, **52**, 272–282.
28. Phillips,K., Dauter,Z., Murchie,A.I., Lilley,D.M. and Luisi,B. (1997) The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution. *J. Mol. Biol.*, **273**, 171–182.
29. Aboul-ela,F., Murchie,A.I., Norman,D.G. and Lilley,D.M. (1994) Solution structure of a parallel-stranded tetraplex formed by d(TG4T) in

the presence of sodium ions by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.*, **243**, 458–471.
30. Aboul-ela,F., Murchie,A.I. and Lilley,D.M. (1992) NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG4T). *Nature*, **360**, 280–282.
31. Kelly,J.A., Feigon,J. and Yeates,T.O. (1996) Reconciliation of the X-ray and NMR structures of the thrombin-binding aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, **256**, 417–422.
32. Macaya,R.F., Schultze,P., Smith,F.W., Roe,J.A. and Feigon,J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl Acad. Sci. USA*, **90**, 3745–3749.
33. Padmanabhan,K., Padmanabhan,K.P., Ferrara,J.D., Sadler,J.E. and Tulinsky,A. (1993) The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *J. Biol. Chem.*, **268**, 17651–17654.
34. Schultze,P., Hud,N.V., Smith,F.W. and Feigon,J. (1999) The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide d(G(4)T(4)G(4)). *Nucleic Acids Res.*, **27**, 3018–3028.
35. Kang,C., Zhang,X., Ratliff,R., Moyzis,R. and Rich,A. (1992) Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature*, **356**, 126–131.
36. Smith,F.W. and Feigon,J. (1993) Strand orientation in the DNA quadruplex formed from the *Oxytricha* telomere repeat oligonucleotide d(G4T4G4) in solution. *Biochemistry*, **32**, 8682–8692.
37. Leonard,G.A., Zhang,S., Peterson,M.R., Harrop,S.J., Helliwell,J.R., Cruse,W.B., d'Estaintot,B.L., Kennard,O., Brown,T. and Hunter,W.N. (1995) Self-association of a DNA loop creates a quadruplex: crystal structure of d(GCATGCT) at 1.8 Å resolution. *Structure*, **3**, 335–340.
38. Kettani,A., Bouaziz,S., Wang,W., Jones,R.A. and Patel,D.J. (1997) *Bombyx mori* single repeat telomeric DNA sequence forms a G-quadruplex capped by base triads. *Nature Struct. Biol.*, **4**, 382–389.
39. Scaria,P.V., Shire,S.J. and Shafer,R.H. (1992) Quadruplex structure of d(G3T4G3) stabilized by K^+ or Na^+ is an asymmetric hairpin dimer. *Proc. Natl Acad. Sci. USA*, **89**, 10336–10340.
40. Keniry,M.A., Strahan,G.D., Owen,E.A. and Shafer,R.H. (1995) Solution structure of the Na^+ form of the dimeric guanine quadruplex [d(G3T4G3)]2. *Eur. J. Biochem.*, **233**, 631–643.
41. Hud,N.V., Smith,F.W., Anet,F.A. and Feigon,J. (1996) The selectivity for K^+ versus Na^+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by ^{1}H NMR. *Biochemistry*, **35**, 15383–15390.
42. Wang,Y. and Patel,D.J. (1995) Solution structure of the *Oxytricha* telomeric repeat d[G4(T4G4)3] G- tetraplex. *J. Mol. Biol.*, **251**, 76–94.
43. Smith,F.W. and Feigon,J. (1992) Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature*, **356**, 164–168.
44. Wang,Y. and Patel,D.J. (1994) Solution structure of the *Tetrahymena* telomeric repeat d(T2G4)4 G-tetraplex. *Structure*, **2**, 1141–1156.
45. Jing,N. and Hogan,M.E. (1998) Structure–activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J. Biol. Chem.*, **273**, 34992–34999.
46. Marathias,V.M. and Bolton,P.H. (1999) Determinants of DNA quadruplex structural type: sequence and potassium binding. *Biochemistry*, **38**, 4355–4364.
47. Morgan,D.M. (1998) Tetrazolium (MTT) assay for cellular viability and activity. *Methods Mol. Biol.*, **79**, 179–183.
48. Mergny,J.L., Phan,A.T. and Lacroix,L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74–78.
49. Jing,N., Rando,R.F., Pommier,Y. and Hogan,M.E. (1997) Ion selective folding of loop domains in a potent anti-HIV oligonucleotide. *Biochemistry*, **36**, 12498–12505.
50. Hardin,C.C., Perry,A.G. and White,K. (2001) Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers*, **56**, 147–194.
51. Rodger,A. and Nordén,B. (1997) *Circular Dichroism and Linear Dichroism. Oxford Chemistry Masters; 1*. Oxford University Press, Oxford.
52. Balagurumoorthy,P. and Brahmachari,S.K. (1994) Structure and stability of human telomeric sequence. *J. Biol. Chem.*, **269**, 21858–21869.
53. Hardin,C.C., Henderson,E., Watson,T. and Prosser,J.K. (1991) Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates. *Biochemistry*, **30**, 4460–4472.
54. Balagurumoorthy,P., Brahmachari,S.K., Mohanty,D., Bansal,M. and Sasisekharan,V. (1992) Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res.*, **20**, 4061–4067.
55. Parkinson,G.N., Lee,M.P. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
56. Matsugami,A., Ouhashi,K., Kanagawa,M., Liu,H., Kanagawa,S., Uesugi,S. and Katahira,M. (2001) An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. *J. Mol. Biol.*, **313**, 255–269.
57. Porumb,H., Monnot,M. and Fermandjian,S. (2002) Circular dichroism signatures of features simultaneously present in structured guanine-rich oligonucleotides: a combined spectroscopic and electrophoretic approach. *Electrophoresis*, **23**, 1013–1020.
58. Jing,N., Marchand,C., Liu,J., Mitra,R., Hogan,M.E. and Pommier,Y. (2000) Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides *in vitro*. *J. Biol. Chem.*, **275**, 21460–21467.
59. Laughlan,G., Murchie,A.I., Norman,D.G., Moore,M.H., Moody,P.C., Lilley,D.M. and Luisi,B. (1994) The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science*, **265**, 520–524.
60. Salisbury,S.A., Wilson,S.E., Powell,H.R., Kennard,O., Lubini,P., Sheldrick,G.M., Escaja,N., Alazzouzi,E., Grandas,A. and Pedroso,E. (1997) The bi-loop, a new general four-stranded DNA motif. *Proc. Natl Acad. Sci. USA*, **94**, 5515–5518.
61. Strahan,G.D., Keniry,M.A. and Shafer,R.H. (1998) NMR structure refinement and dynamics of the K^+-[d(G3T4G3)]2 quadruplex via particle mesh Ewald molecular dynamics simulations. *Biophys. J.*, **75**, 968–981.
62. Smith,F.W., Schultze,P. and Feigon,J. (1995) Solution structures of unimolecular quadruplexes formed by oligonucleotides containing *Oxytricha* telomere repeats. *Structure*, **3**, 997–1008.

DNA Interactions of Monofunctional Organometallic Ruthenium(II) Antitumor Complexes in Cell-free Media†

Olga Novakova,‡ Haimei Chen,§ Oldrich Vrana,‡ Alison Rodger,|| Peter J. Sadler,*,§ and Viktor Brabec*,‡

Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic, School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, United Kingdom, and Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

Received May 30, 2003; Revised Manuscript Received August 5, 2003 Publication 82

ABSTRACT: Modifications of natural DNA in a cell-free medium by antitumor monodentate Ru(II) arene compounds of the general formula $[(\eta^6\text{-arene})Ru(en)Cl]^+$ (arene = biphenyl, dihydroanthracene, tetrahydroanthracene, *p*-cymene, or benzene; en = ethylenediamine) were studied by atomic absorption, melting behavior, transcription mapping, circular and linear dichroism, plasmid unwinding, competitive ethidium displacement, and differential pulse polarography. The results indicate that these complexes bind preferentially to guanine residues in double-helical DNA. The data are consistent with DNA binding of the complexes containing biphenyl, dihydroanthracene, or tetrahydroanthracene ligands that involves combined coordination to G N7 and noncovalent, hydrophobic interactions between the arene ligand and DNA, which may include arene intercalation and minor groove binding. In contrast, the single hydrocarbon rings in the *p*-cymene and benzene ruthenium complexes cannot interact with double-helical DNA by intercalation. Interestingly, the adducts of the complex containing *p*-cymene ligand, which has methyl and isopropyl substituents, distort the conformation and thermally destabilize double-helical DNA distinctly more than the adducts of the three multiring ruthenium arene compounds. It has been suggested that the different character of conformational alterations induced in DNA, and the resulting thermal destabilization, may affect differently further "downstream" effects of damaged DNA and consequently may result in different biological effects of this new class of metal-based antitumor compounds. The results point to a unique profile of DNA binding for Ru(II) arene compounds, suggesting that a search for new anticancer compounds based on this class of complexes may also lead to an altered profile of biological activity in comparison with that of metal-based antitumor drugs already used in the clinic or currently on clinical trials.

Platinum coordination compounds are widely used as antitumor drugs. The first platinum antitumor drugs introduced in the clinic were *cis*-diamminedichloroplatinum(II)[1] (cisplatin) and its somewhat less toxic analogue carboplatin [*cis*-diammine-1,1-cyclobutanedicarboxylatoplatinum(II)]. Both compounds show the same spectrum of antitumor activity. The clinical efficacy of these anticancer drugs is diminished by intrinsic and acquired tumor resistance. Owing to these limitations, there is an intense effort to design new transition-metal-based compounds that are capable of overcoming problems associated with cisplatin and carboplatin chemotherapy while delivering the therapeutic effect.

Cisplatin and carboplatin target cellular DNA, forming covalent adducts (*1*). The most abundant adducts formed by cisplatin in linear DNA are 1,2-d(GpG) (∼65%) or 1,2-d(ApG) (25%) intrastrand and 1,2-d(GG) (∼6%) interstrand cross-links (*2*−*4*). In the years following the introduction of cisplatin, the design of new platinum antitumor drugs concentrated mainly on direct cisplatin analogues which adhered to the set of structure−activity relationships established as early as 1973 (*5*, *6*). More recently, there have been efforts to design rationally unconventional platinum complexes that violate these original structure−activity relationships, such as polynuclear platinum complexes or analogues of the clinically ineffective trans isomer of cisplatin (transplatin) (*7*−*9*). There have also been efforts directed at the

† Supported by the Grant Agency of the Czech Republic (Grant Nos. 305/02/1552, 305/01/0418), the Grant Agency of the Academy of Sciences of the Czech Republic (Grant No. A5004101), The Wellcome Trust (U.K.), the Committee of Vice-Chancellors and Vice-Principals (Overseas Research Scholarship to HC), and the Edinburgh Technology Fund. The authors acknowledge that this research is also a part of the European Cooperation in the field of Scientific and Technical Research network (COST projects D20/0003/00 and D20/0005/01).

* Corresponding authors. VB: telephone, 420-541517148; fax, 420-541240499; e-mail, brabec@ibp.cz. PJS: telephone, +44-131-6504729; fax, +44-131-6506452; e-mail, P.J.Sadler@ed.ac.uk.

‡ Institute of Biophysics, Academy of Sciences of the Czech Republic.

§ School of Chemistry, University of Edinburgh.

|| Department of Chemistry, University of Warwick.

[1] Abbreviations: BEN, benzene; BIP, biphenyl; CD, circular dichroism; cisplatin, *cis*-diamminedichloroplatinum(II); CT, calf thymus; CYM, *p*-cymene; DHA, dihydroanthracene; DPP, differential pulse polarography; en, ethylenediamine; EtBr, ethidium bromide; FAAS, flameless atomic absorption spectrophotometry; LD, linear dichroism; [Pt(dien)Cl]Cl, chlorodiethylenetriamineplatinum(II) chloride; dien = diethylenetriamine; poly(dG-dC), polymer of alternating dG and dC residues; poly(dA-dT), polymer of alternating dA and dT residues; r_b, the number of metal atoms bound per nucleotide residue; r_i, the molar ratio of free metal complex to nucleotide-phosphates at the onset of incubation with DNA.; THA, tetrahydroanthracene; t_m, DNA melting temperature; transplatin, *trans*-diamminedichloroplatinum(II).

10.1021/bi034933u CCC: $25.00
Published on Web 09/06/2003

design of other transition-metal antitumor agents. Possible advantages in using transition-metal ions other than Pt(II) include the availability of additional coordination sites in octahedral complexes and the altered shape of the complex, alterations in ligand affinity and substitution kinetics, changes in oxidation state, and photodynamic approaches to therapy. In the design of these new drugs, ruthenium complexes have attracted much interest (*10*−*12*).

Organometallic ruthenium(II) arene complexes of the type [(η^6-arene)Ru(II)(en)Cl][PF_6] (en = ethylenediamine) constitute a relatively new group of anticancer compounds (*13*, *14*). These are pseudo-octahedral "piano-stool" complexes with the arene ring occupying three coordination sites and two diamine nitrogens and the halide ligand occupying the remaining three sites. Importantly, the arene ligand is strongly bound and stabilizes ruthenium(II): the complexes do not readily undergo oxidation to ruthenium(III). Moreover, most of the compounds are ionic and have a reasonable aqueous solubility. These monodentate Ru(II) arene complexes have been considered novel anticancer agents with a mechanism of action different from that of the ruthenium(III) complex (ImH)[*trans*-Ru(III)Cl_4Im(Me_2SO)] {Im = imidazole, NAMI-A}, which is currently on clinical trial (*11*, *15*).

Broadening the chemotherapeutic arsenal depends on understanding existing agents with a view toward developing new modes of attack. Interestingly, a direct correlation between cytotoxicity and DNA binding has already been observed for several ruthenium compounds in cell cultures (*16*). Also consistent with DNA binding in vivo, a number of ruthenium compounds inhibit DNA replication, exhibit mutagenic activity, induce the SOS repair mechanism, bind to nuclear DNA, and reduce RNA synthesis (*16*). Similarly, several antitumor ruthenium compounds have also been shown (*17*−*21*) to bind to DNA and inhibit DNA replication in vitro, but the pharmacological target for ruthenium compounds and the mechanism underlying its biological effects are not known.

The (η^6-arene)Ru(II) bonds are inert toward hydrolysis, but the monofunctional complexes [(η^6-arene)Ru(II)(en)(Cl)]$^+$ readily lose their chloride ligand and transform into the corresponding, more reactive, aquated species (*22*). It has also been shown (*14*) that the complex [(η^6-*p*-cymene)Ru(II)(en)(Cl)]$^+$ binds to a short, single-stranded deoxyribooligonucleotide (14-mer), forming monofunctional adducts, and that the preferential sites of ruthenation in this oligomer are guanine residues. To address further fundamental questions about DNA binding modes of ruthenium(II) arene antitumor compounds, the experiments described in the present paper were carried out. More specifically, the interactions of polymeric B-DNAs with [(η^6-arene)Ru(II)(en)(Cl)]$^+$ complexes [where arene = biphenyl (BIP), dihydroanthracene (DHA), tetrahydroanthracene (THA), *p*-cymene (CYM), or benzene (BEN)] in cell-free media were investigated by various biochemical and biophysical methods with the goal to contribute to understanding their biological effects and to help establish structure−pharmacological relationships for this class of ruthenium anticancer compounds.

MATERIALS AND METHODS

Starting Materials. [(η^6-arene)Ru(II)(en)(Cl)]PF_6 complexes (Figure 1) were prepared as described previously (*23*).

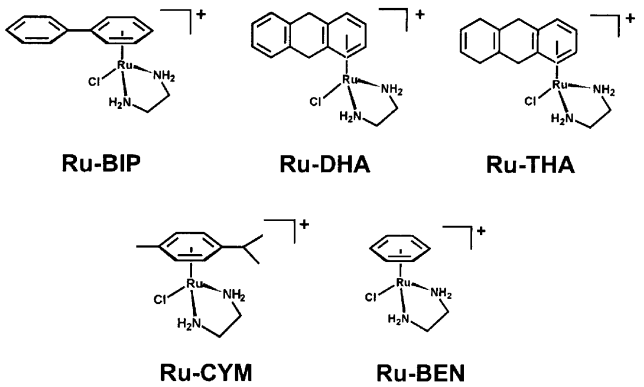


FIGURE 1: Structures of Ru(II) arene complexes used in this work.

Cisplatin, *trans*-diamminedichloroplatinum(II) (transplatin), and chlorodiethylenetriamineplatinum(II) chloride {[Pt(dien)Cl]Cl} were synthesized and characterized at Lachema (Brno, Czech Republic). The stock solutions of the ruthenium and platinum complexes (5×10^{-4} M in H_2O) were prepared in the dark at 25 °C and stored for at least 7 days before they were used. Aquation of these Ru(II) arene complexes is relatively rapid in water (*22*), and as for cisplatin, the aqua adducts are more reactive than the chloro complexes. Calf thymus (CT) DNA (42% G + C, mean molecular mass ~ 2×10^7) was also prepared and characterized as described previously (*24*, *25*). Poly(dG-dC) and poly(dA-dT) were obtained from Amersham Pharmacia-Biotech (Piscataway, NJ) and were used without further purification.

Plasmid pSP73KB [2455 bp (*26*)] was isolated according to standard procedures and banded twice in CsCl/EtBr equilibrium density gradients. Restriction endonucleases were purchased from New England Biolabs (Beverly, MA). Riboprobe Gemini System II for transcription mapping containing SP6 and T7 RNA polymerases was purchased from Promega (Madison, WI). Ethidium bromide (EtBr) and agarose were from Merck KgaA (Darmstadt, Germany). The radioactive products were from Amersham (Arlington Heights, IL).

Metalation Reactions. CT DNA and plasmid DNAs were incubated with ruthenium or platinum complex in 10 mM $NaClO_4$ (pH ~ 6) at 37 °C for 48 h in the dark, if not stated otherwise. The number of atoms of the metal bound per nucleotide residue (r_b values) was determined by flameless atomic absorption spectrophotometry (FAAS) (*21*).

DNA Transcription by RNA Polymerase in Vitro. Transcription of the (*Nde*I/*Hpa*I) restriction fragment of pSP73KB DNA with SP6 or T7 RNA polymerase and electrophoretic analysis of transcripts were performed according to the protocols recommended by Promega (Promega Protocols and Applications, 43−46 (1989/90)) and previously described in detail (*26*, *27*).

Unwinding of Negatively Supercoiled DNA. Unwinding of closed circular supercoiled pSP73KB plasmid DNA was assayed by an agarose gel mobility shift assay (*28*). The unwinding angle, Φ, induced per metal−DNA adduct was calculated upon the determination of the r_b value at which the complete transformation of the supercoiled form to the relaxed form of the plasmid was attained. Samples of pSP73 plasmid were incubated with Ru(II) arene or cisplatin in 10 mM $NaClO_4$ at 37 °C in the dark for 48 h. All samples were precipitated by ethanol and redissolved in TAE buffer (0.04

M Tris-acetate + 1 mM EDTA, pH 7.0). An aliquot of the precipitated sample was subjected to electrophoresis on 1% agarose gels running at 25 °C in the dark with TAE buffer with a voltage set at 30 V. The gels were then stained with EtBr, followed by photography on Polaroid 667 film with transilluminator. The other aliquot was used for the determination of r_b values by FAAS.

Circular Dichroism (CD) and Linear Dichroism (LD). If not stated otherwise, CD and LD spectra of DNA modified by the ruthenium complexes were recorded at 25 °C in 10 mM $NaClO_4$ on JASCO J-720 and adapted J-715 spectropolarimeters.

Differential Pulse Polarography. Differential pulse polarographic (DPP) curves of DNA were measured after nonmodified DNA or DNA modified by either Ru(II) arene compound used in the present work was redissolved in a medium consisting of 0.3 M ammonium formate plus 0.05 M phosphate (Na_2HPO_4/NaH_2PO_4) buffer, pH 6.8. DPP curves were recorded with the aid of an EG&C PARC electrochemical analyzer, model 384B at 25 °C using the following apparatus settings: voltage scan rate of 2 mV/s, pulse amplitude of 5 mV, drop time of 1.0 s. The potentials are relative to the saturated calomel reference electrode (SCE).

Fluorescence Measurements. Fluorescence measurements in the presence of EtBr of CT DNA modified by ruthenium arene complexes were performed at an excitation wavelength of 546 nm, and the emitted fluorescence was analyzed at 590 nm. The fluorescence intensity was measured at 25 °C in 0.4 M NaCl to avoid secondary binding of EtBr to DNA (*29*, *30*). The concentrations were 0.01 mg/mL for DNA and 0.04 mg/mL for EtBr, which corresponded to the saturation of all intercalation sites of EtBr in DNA (*29*, *30*). These measurements were performed on a Shimadzu RF 40 spectrofluorophotometer using a 1 cm quartz cell.

DNA Melting. The melting curves of DNA were recorded by measuring the absorbance at 260 nm. The melting curves of unruthenated or ruthenated DNA were recorded after Tris-HCl/EDTA buffer and $NaClO_4$ were added so that the resulting media contained 0.01−0.2 M $NaClO_4$ with 1 mM Tris-HCl/0.1 mM EDTA, pH 7.4. The value of the melting temperature (t_m) was determined as the temperature corresponding to a maximum on the first-derivative profile of the melting curves. The t_m values were thus determined with an accuracy of ±0.3 °C.

Other Methods. Absorption spectra were measured with Beckmann DU-7400 and Cary 1E spectrophotometers. FAAS measurements were carried out on a Unicam 939 AA spectrometer with a graphite furnace. For FAAS analysis, DNA was precipitated with ethanol and dissolved in 0.1 M HCl.

RESULTS

DNA Binding. Solutions of double-helical CT DNA at a concentration of 0.1 mg/mL were incubated with Ru(II) arene complexes at an r_i (molar ratio of free ruthenium complex to nucleotide phosphates at the onset of incubation with DNA) value of 0.1 in 10 mM $NaClO_4$ at 37 °C. At various time intervals, an aliquot of the reaction mixture was withdrawn, quickly cooled on an ice bath, and precipitated by ethanol, and the content of ruthenium in the supernatant of these samples was determined by FAAS. Figure 2 shows a plot of r_b (the number of atoms of metal bound per nucleotide residue) against the time of DNA incubation with Ru−BIP (r_i = 0.1). The amount of ruthenium bound per DNA nucleotide phosphate (r_b) increased with time. After ~3 h, approximately 90% of the molecules of the Ru−BIP present in the reaction mixture were bound to DNA. In these binding reactions, the time at which the binding reached 50% ($t_{50\%}$) was ~10, 15, and 10 min and 3.5 h for the compounds Ru−BIP, Ru−DHA, Ru−THA, and Ru−CYM, respectively. The value of $t_{50\%}$ for the reaction of cisplatin with DNA under conditions identical to those specified in Figure 2 was ~2 h (*31*).

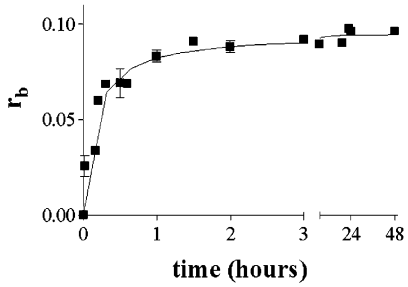


FIGURE 2: Kinetics of the binding of the Ru(II) arene compound Ru−BIP to calf thymus DNA. Medium: 10 mM $NaClO_4$ at 37 °C. The concentration of DNA was 0.1 mg/mL; r_i = 0.1. Data measured in triplicate varied on average ±3% from their mean.

The binding of Ru(II) arene compounds to CT DNA was also quantified in two other ways. Aliquots of the reaction mixture withdrawn at various time intervals were quickly cooled on an ice bath and then exhaustively dialyzed against 10 mM $NaClO_4$ at 4 °C or filtered using Sephadex G50 to remove free (unbound) ruthenium compound. The content of ruthenium in these DNA samples was determined by FAAS. Results identical to those obtained using the assay based on DNA precipitation by ethanol were obtained.

In further experiments, CT DNA was also incubated with the Ru(II) arene complexes at r_i = 0.2 and essentially a similar rate of binding was observed as for the reaction at r_i= 0.1. When Ru complexes were added to a higher level (r_i = 0.25), the DNA precipitated immediately, probably due to electrostatic interactions between the positively charged Ru(II) arene complexes and DNA.

The binding experiments carried out in this work indicated that modification reactions resulted in the irreversible coordination of the Ru(II) arene complexes to polymeric double-helical DNA, which thus facilitated sample analysis. Hence, it was possible to prepare easily and precisely samples of DNA modified by the ruthenium complex at a preselected value of r_b. Thus, except where stated, samples of DNA modified by Ru(II) arene compounds and analyzed further by biophysical or biochemical methods were prepared in 10 mM $NaClO_4$ at 37 °C. After 24 h of the reaction of DNA with the complex, the samples were precipitated in ethanol and dissolved in the medium necessary for a particular analysis, and the r_b value in an aliquot of this sample was checked by FAAS. In this way, most of the analyses described in the present paper were performed in the absence of unbound (free) Ru(II) arene complex.

In Vitro Transcription of DNA Containing Ru(II) Arene Adducts. In vitro RNA synthesis by RNA polymerases on DNA templates containing several types of bifunctional

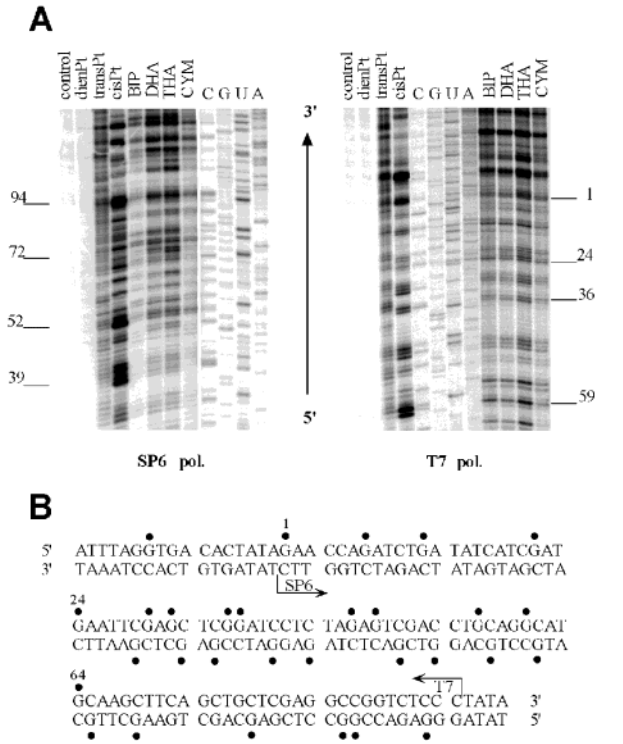


FIGURE 3: Inhibition of RNA synthesis by SP6 and T7 RNA polymerases on the *Nde*I/*Hpa*I fragment of pSP73KB plasmid modified by Ru(II) arene and platinum complexes. (A) Autoradiograms of 6% polyacrylamide/8 M urea sequencing gels showing inhibition of RNA synthesis by SP6 (left) or T7 RNA polymerases (right) on the *Nde*I/*Hpa*I fragment containing adducts of ruthenium or platinum complexes. Lanes: control, unmodified template; BIP, DHA, THA, CYM, cisPt, transPt, and dienPt, the template modified by Ru−BIP, Ru−DHA, Ru−THA, Ru−CYM, cisplatin, transplatin, or [Pt(dien)Cl]Cl at $r_b = 0.01$, respectively; A, U, G, and C, chain terminated marker RNAs. (B) Schematic diagram showing the portion of the sequence used to monitor inhibition of RNA synthesis by ruthenium and platinum complexes. The arrows indicate the start of the SP6 and T7 RNA polymerase, which used as template the bottom or upper strand of the NdeI/HpaI fragment of pSP73KB DNA, respectively. (●) major stop signals (from Figure 3A) for DNA modified by Ru−BIP. The numbers correspond to the nucleotide numbering in the sequence map of pSP73KB plasmid.

adducts of platinum complexes can be prematurely terminated at the level or in the proximity of adducts (*26*, *27*). Importantly, monofunctional DNA adducts of several platinum complexes are unable to terminate RNA synthesis (*26*, *27*, *32*).

Cutting of pSP73KB DNA by *Nde*I and *Hpa*I restriction endonucleases yielded a 212-bp fragment (*26*, *27*) (a substantial part of its nucleotide sequence is shown in Figure 3B). This fragment contained SP6 or T7 RNA polymerase promotors [in both strands close to their 3′-ends (Figure 3B)]. The experiments were carried out using this linear DNA fragment modified by Ru(II) arene complexes, cisplatin, transplatin, or [Pt(dien)Cl]Cl at $r_b = 0.01$, for RNA synthesis by SP6 or T7 RNA polymerase (Figure 3A, lanes BIP, DHA, THA, CYM, cisDDP, transDDP, or dienPt, respectively). RNA synthesis on the fragment modified by the ruthenium and bifunctional platinum complexes yielded fragments of defined sizes, which indicates that RNA synthesis on these templates was prematurely terminated. The major stop sites produced by ruthenium compounds were identical for all four Ru(II) arene compounds Ru−BIP, Ru−DHA, Ru−THA, and Ru−CYM (Ru−BEN was not studied in this assay) and were mainly at guanine residues. The corresponding bands on the autoradiogram were of similar intensity for the compounds Ru−BIP, Ru−DHA, and Ru−THA, whereas the bands produced by Ru−CYM were markedly less intense. For comparative purposes, the inhibition of RNA synthesis by DNA adducts of cisplatin, transplatin, and monofunctional [Pt(dien)Cl]Cl is also shown (Figure 3A, lanes cisPt, transPt, and dienPt) and demonstrates more termination for cisplatin but at the same termination sites as those for the Ru(II) arenes, different termination sites than those for transplatin, and no termination of RNA synthesis by monofunctional [Pt(dien)Cl]Cl. The sequence analysis reveals that the major bands resulting from termination of RNA synthesis by the adducts of cisplatin and Ru(II) arene compounds preferentially appear one or a half nucleotide preceding G sites and to a considerably less extent preceding A sites (in AGAG or AGGAG sequences). Taken together, Ru(II) arene compounds exhibit a base sequence selectivity similar to that of cisplatin. Nevertheless, the efficiency of the adducts of Ru(II) arene complexes to terminate RNA synthesis in vitro is in general reduced relative to that of cisplatin. Furthermore, the efficiency of the Ru−CYM complex was noticeably lower than that of the other three ruthenium compounds examined in this assay.

Circular and Linear Dichroism Spectroscopy. A sizable CD spectrum can be induced into the absorbance bands of a number of low-molecular-mass molecules upon their binding to double-helical DNA (*33*). Ideally, ligand absorbance bands that do not overlap with DNA bands (below 300 nm) are used to probe the interaction. The Ru(II) arene compounds investigated here are achiral (not optically active) when free in solution. Binding of the Ru(II) arene complexes to DNA was indicated by the induction of a sizable CD spectrum for the Ru(II)−arene absorption bands [e.g. Ru(II) to arene π^* charge-transfer transitions at 375 nm] in the presence of double-helical DNA.

CD spectra for CT DNA in the absence and in the presence of increasing amounts of Ru(II) arene compounds bound to DNA are compared in Figure 4 (panels A−D). Binding of the compounds Ru−BIP, Ru−DHA, and Ru−THA to CT DNA results in the appearance of a positive CD band centered around 370−380 nm, and its intensity increased with the level of DNA ruthenation. Interestingly, the induced CD signal of Ru−BIP was the largest, whereas the binding of Ru−CYM to double-helical DNA resulted in no induced CD band in this region (Figure 4D). The different intensities of the induced CD bands produced by the compounds Ru−BIP, Ru−DHA, and Ru−THA are, at least partly, due to the different extinction coefficients (ca. 380, 280, and 150 M^{-1} cm^{-1}, respectively) of the free compounds at the wavelengths corresponding to this CD band.

Upon binding of Ru(II) arene compounds to CT DNA, the approximately conservative CD spectrum normally found for B-DNA at wavelengths below 300 nm (Figure 4A−D) is also changed. As a function of r_b, there was a marked increase in the intensity of the positive band around 280 nm when DNA was modified by the compounds Ru−BIP, Ru−DHA, and Ru−THA, accompanied by a decrease in the intensity of the negative band at 245 nm (Figure 4A−C). This induced CD signal is due to changes in both the intrinsic DNA CD and the ligand-induced CD. On the other hand, when DNA was modified by Ru−CYM, no such increase

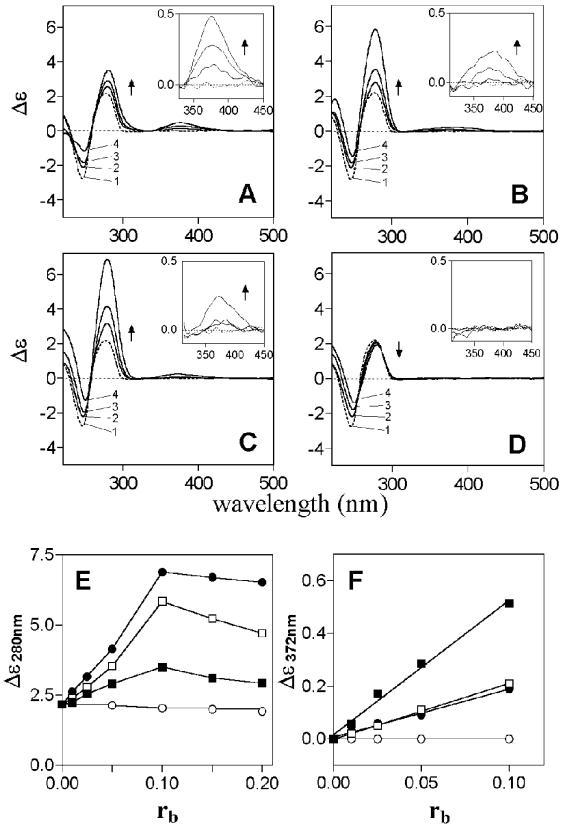


FIGURE 4: CD spectroscopy of calf thymus DNA modified by Ru(II) arene compounds. The spectra were recorded for DNA at the concentration 30 μg/mL in 10 mM $NaClO_4$. (A−D) CD spectra of DNA modified by Ru−BIP, Ru−DHA, Ru−THA, or Ru−CYM, respectively. Curves in parts A−C (from the bottom to the top at ∼250, ∼275, and 375 nm): 1 (- - -), control (nonmodified) DNA; 2, $r_b = 0.025$; 3, $r_b = 0.05$; 4, $r_b = 0.1$. Curves in part D (at ∼250 nm, the order of appearance of the curves is the same as that given for Ru−BIP, Ru−DHA, and Ru−THA in parts A−C, whereas it is reversed at ∼275 nm): 1 (- - -), control (nonmodified) DNA; 2, $r_b = 0.05$; 3, $r_b = 0.1$; 4, $r_b = 0.2$. Inset in parts A−D: a part of the CD spectrum (at 280−450 nm) recorded at a higher sensitivity of the CD instrument. (E) Dependence of the maximum ellipticity of the positive CD band at around 280 nm on r_b: (■) Ru−BIP; (□) Ru−DHA; (●) Ru−THA; (○) Ru−CYM. (F) Dependence of the maximum ellipticity of the positive CD band at around ∼370 nm on r_b: (■) Ru−BIP; (□) Ru−DHA; (●) Ru−THA; (○) Ru−CYM. Data points measured in duplicate varied on average ±1% from their mean.

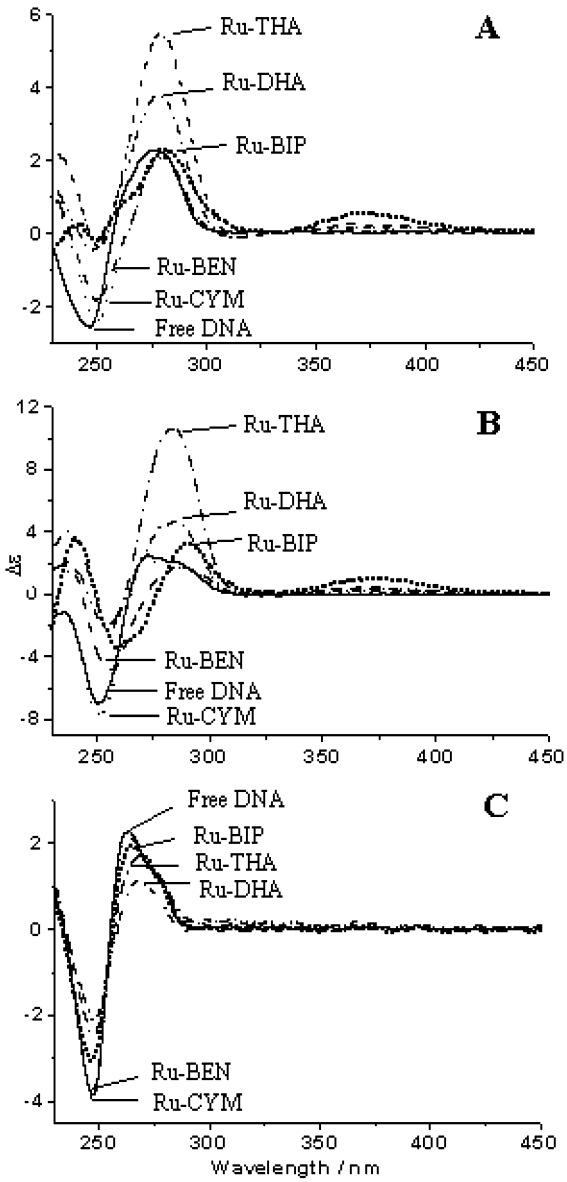


FIGURE 5: Comparisons of the effects of different arene ligands on the CD spectra of (A) CT DNA, (B) poly(dG-dC), and (C) poly(dA-dT) in 10 mM $NaClO_4$ in the presence of various Ru(arene) complexes at $r_i = 0.2$. The spectra were recorded at 25 °C after samples had been incubated for 24 h at 37 °C.

in the intensity of the positive band around 280 nm was seen, but instead there was a small decrease.

We also recorded CD spectra of CT DNA and the synthetic double-stranded polynucleotide complexes poly(dG-dC) and poly(dA-dT) in the presence of these four Ru complexes and also the benzene complex Ru(BEN) at $r_i = 0.2$. An induced CD band at 350−410 nm was observed not only for interaction of Ru−BIP, Ru−THA, and Ru−DHA with CT DNA but also for interaction with poly(dG-dC) (Figure 5A and B). For poly(dG-dC), as for CT DNA, no induced band in the near UV region was observed for Ru−CYM, or for Ru−BEN (Figure 5A and B). For poly(dA-dT), only the complexes with extended π systems, Ru−BIP, Ru−THA, and Ru−DHA, induced any spectral changes, and these were small and confined to the region 260−280 nm (Figure 5C). These small changes may be due to weak hydrophobic interactions between the arene and the DNA bases. Neither Ru−CYM nor Ru−BEN induced any spectral changes in poly(dA-dT).

The flow linear dichroism data for CT DNA (Figure S1A) resemble those of poly(dG-dC) DNA (Figure S1B) and show that the binding of all Ru complexes causes bending of the DNA. The three complexes Ru−THA, Ru−DHA, and Ru−BIP cause a significant red shift (ca. 10 nm) of the main DNA band near 260 nm, whereas the two complexes Ru−BEN and Ru−CYM cause no shift. The wavelength shifts in the region of DNA absorption for Ru−THA, Ru−DHA, and Ru−BIP are consistent with intercalation of the arene ligands, but the bending precludes full intercalation which would rigidify the DNA and thus increase the LD. Ru−BEN and Ru−CYM complexes rigidify poly(dA-dT) DNA with no 260 nm band shift (Figure S1C); the other three complexes, Ru−DHA, Ru−THA, and Ru−BIP, cause significant bending while retaining sufficient orientation to produce an LD signal for these compounds. Ru−DHA

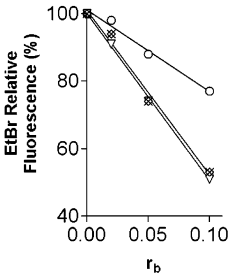


FIGURE 6: Dependences of the EtBr fluorescence on r_b for DNA modified by Ru(arene) complexes in 10 mM $NaClO_4$ at 37 °C for 24 h: (×) Ru−BIP; (◇) Ru−DHA; (▽) Ru−THA; (○) Ru−CYM. Data points measured in triplicate varied on average ±2% from their mean.

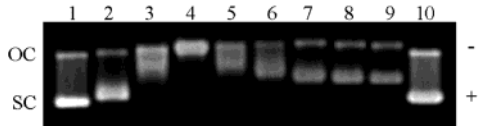


FIGURE 7: Unwinding of supercoiled pSP73KB plasmid DNA by the compound Ru−BIP. The plasmid was incubated with the ruthenium complex for 24 h at 37 °C. Lanes: 1 and 10, control, nonmodified DNA ($r_b = 0$); 2, $r_b = 0.03$; 3, $r_b = 0.06$; 4, $r_b = 0.08$; 5, $r_b = 0.10$; 6, $r_b = 0.11$; 7, $r_b = 0.13$; 8, $r_b = 0.16$; 9, $r_b = 0.17$. The top bands correspond to the form of nicked plasmid, and the bottom bands, to the closed, negatively supercoiled plasmid.

binding induces the most bending. The LD observed for the binding to poly(dA-dT) is consistent with a mode involving the aromatic ligands of Ru−DHA, Ru−THA, and Ru−BIP inserted into the minor groove, giving a positive signal for the transitions in the 260−280 nm ligand y-polarized region and probably a negative one for the 240−255 nm z-polarized region (according to the film LD assignments, Figure S2).

Characterization of DNA Adducts by EtBr Fluorescence. EtBr as a fluorescent probe can be used to distinguish intercalating and nonintercalating ligands (*34*). Binding of EtBr to DNA by intercalation is blocked in a stoichiometric manner by formation of a wide spectrum of DNA-binding ligands including intercalators. On the other hand, modification of DNA by monofunctional nonintercalative ligands, such as [Pt(dien)Cl]Cl, results in only a slight decrease of EtBr fluorescence intensity as compared with that for the complex of nonmodified DNA with EtBr. Competitive binding of other intercalators leads to a loss of fluorescence because of depletion of the DNA−EtBr complex (free EtBr is poorly fluorescent).

Double-helical DNA was modified by Ru arene compounds for 24 h. The levels of the modification corresponded to the values of r_b in the range between 0 and 0.1. Modification of DNA by Ru−BIP, Ru−DHA, and Ru−THA resulted in a marked decrease of EtBr fluorescence (Figure 6). In contrast, the decrease of the fluorescence intensity by the adducts of Ru−CYM was only very small and similar to that induced by the adducts of [PtCl(dien)]Cl (*35*).

DNA Unwinding. Electrophoresis in native agarose gels was used to determine the unwinding induced in negatively supercoiled pSP73KB plasmid DNA by monitoring the degree of supercoiling (*28*) (Figure 7). A compound that unwinds the DNA duplex reduces the number of supercoils in closed circular DNA, which in turn causes a decrease in the rate of migration through the agarose gel.

Figure 7 shows an electrophoresis gel from experiments in which variable amounts of Ru−BIP were bound to a mixture of relaxed and negatively supercoiled pSP73KB DNA. The mean unwinding angle is given by $\Phi = 18\sigma/r_b(c)$, where σ is the superhelical density and $r_b(c)$ is the value of r_b at which the supercoiled and nicked forms comigrate (*28*). Under the present experimental conditions, σ was calculated to be −0.063 on the basis of data for cisplatin for which the value $\Phi = 13°$ was used to determine $r_b(c)$ (*28*). The $r_b(c)$ values for all ruthenium arene compounds studied, along with the mean unwinding angles calculated in this way, are summarized in Table 1. The unwinding angles were 14 ± 1° per bound Ru−BIP, Ru−DHA, and Ru−THA, and 7 ± 0.5° for Ru−CYM.

Table 1: Unwinding of Supercoiled pSP73KB DNA by Ru(II) Arene Complexes

compd	$r_b(c)$	unwinding angle[a]
Ru−BIP	0.08 ± 0.005	14 ± 1°
Ru−DHA	0.08 ± 0.005	14 ± 1°
Ru−THA	0.08 ± 0.005	14 ± 1°
Ru−CYM	0.16 ± 0.01	7 ± 0.5°

[a] The unwinding angle was calculated as described in the text.

DNA Melting. CT DNA was modified by Ru(II) arene compounds added at various r_b values (0−0.1) in 10 mM $NaClO_4$ at 37 °C for 24 h. The salt concentration was then further adjusted by addition of $NaClO_4$ to values in the range 0.01−0.1 M. The effect on t_m is dependent both on the amount of ruthenium bound and on the salt concentration. At low concentrations of $NaClO_4$ (0.01 M) an increase in t_m was observed for the three compounds Ru−BIP, Ru−DHA, and Ru−THA, and this became more pronounced with increasing r_b values (Figure 8A−C). With increasing ionic strength, the enhancement of t_m (Δt_m) due to the presence of the Ru compounds decreased, and at salt concentrations of 0.2 M, t_m decreased. In contrast, the melting behavior of DNA modified by Ru−CYM resulted in a similar decrease of t_m at all ionic strengths (Figure 8D).

Differential Pulse Polarography. This is a sensitive method for distinguishing between nondenaturational and denaturational conformational alterations in DNA induced by various physical or chemical agents (*36*). The analysis is based on the observation that intact double-helical DNA is polarographically inactive because its reduction sites are involved in hydrogen bonds and are unable to make contact with the working electrode in a manner suitable for electron transfer. Electroreduction of adenine or cytosine residues present in distorted but still double-stranded (nondenatured) regions of DNA is responsible for the appearance of the small DPP peak II (Figure 9A, curve 1). Base residues in these distorted regions become more accessible for electroreduction at the mercury electrode and can yield a small polarographic current. On the other hand, the appearance of a more negative peak III in DPP curves of DNA indicates the presence of single-stranded, denatured regions in the DNA molecule, in which hydrogen bonds between complementary bases have been broken (*36*). Differences in the adsorption properties of double-helical and denatured DNA at the mercury electrode have been suggested to give rise to the different reduction potentials observed for the two DNA conformations. Importantly, less than 1% of the denatured material in the presence of an excess of double-helical DNA can be determined by DPP (*37*).

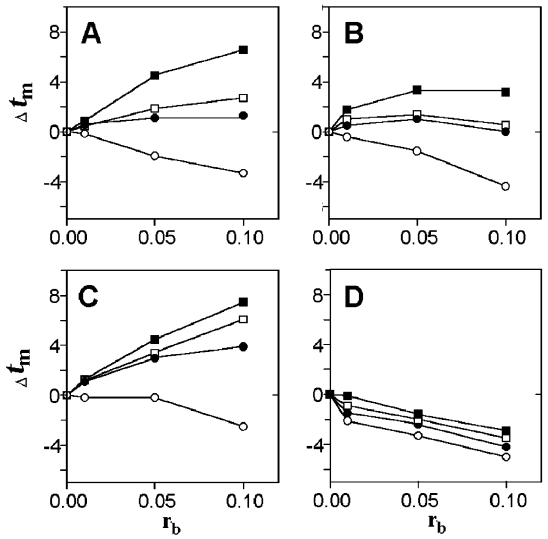


FIGURE 8: Plots showing the dependence of Δt_m values on r_b for calf thymus DNA modified by Ru(II) arene compounds: A, Ru−BIP; B, Ru−DHA; C, Ru−THA; D, Ru−CYM. The melting curves were measured in 0.01 M (■), 0.05 M (□), 0.1 M (●), or 0.2 M (○) $NaClO_4$ plus 1 mM Tris-HCl with 0.1 mM EDTA, pH 7.4. Δt_m is defined as the difference between the t_m values of ruthenated and nonmodified DNAs. Data measured in triplicate varied on average ±2% from their mean.

DPP has already been used to analyze DNA modified by various physical or chemical agents, including platinum and ruthenium compounds with different clinical efficacy (*20*, *38*, *39*). It has been found that DNA globally modified by antitumor cisplatin, its antitumor analogues, or *cis*- or *trans*-$[RuCl_2(Me_2SO)_4]$ at r_b values up to 0.05 yields the more positive DPP peak II, indicating that these antitumor drugs induce nondenaturational conformational changes in DNA (*38*). In contrast, the more negative peak III is evident on DPP curves of DNA globally modified by clinically ineffective transplatin and other inactive platinum(II) complexes, indicating that the clinically ineffective platinum complexes induce denaturational conformational alterations in DNA (*38*, *39*).

DPP analysis also sheds considerable light on the conformational basis for DNA binding of the Ru(II) arene compounds tested in this work. Modification of CT DNA by Ru−BIP, Ru−DHA and Ru−THA at r_b values of 0.0005−0.01 resulted in an increase in intensity of the DPP peak II with increasing levels of the modification (shown for Ru−DHA in Figure 9B). The more negative peak III was not detected even for the DPP curves recorded for DNA modified at the highest r_b value used in our experiments (0.01). As regards the DPP analysis of DNA modified by Ru−THA, peak II was markedly distorted by a current corresponding to the higher background electrolyte discharge, so that the measurement of the height of this peak II at r_b values of 0.001 and higher was impossible. On the other hand, it was clear even from these distorted DPP curves that no peak III was present even at $r_b = 0.01$. It could be argued that the absence of peak III on the DPP curves recorded for the samples of DNA modified by Ru−BIP, Ru−DHA and Ru−THA at relatively high r_b values (~0.02) could be due to an increase in the slope of the part of the DPP curve corresponding to the background electrolyte discharge (Figure 9B, curve 4). The fact that peak III was not buried under the background electrolyte discharge curve was verified using

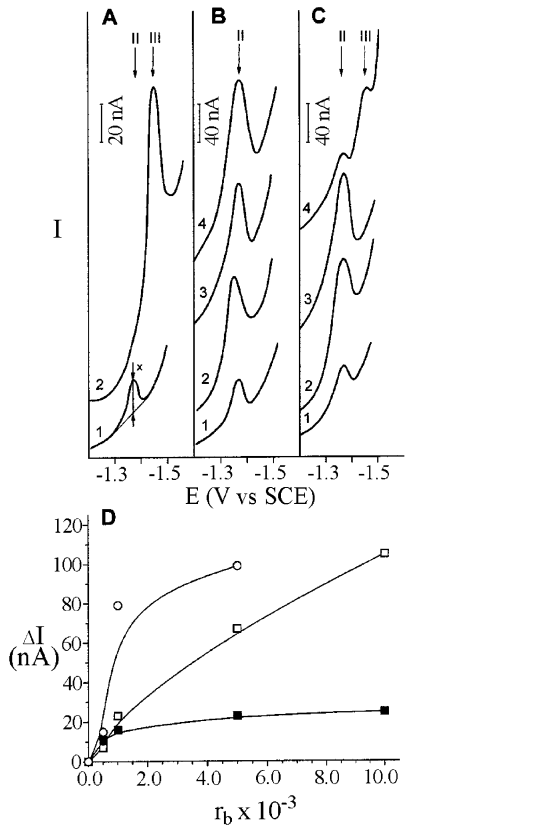


FIGURE 9: Differential pulse polarographic analysis of CT DNA modified by Ru(II) arene compounds. Double-helical DNA at a concentration of 0.32 mg/mL in 0.3 M ammonium formate with 0.01 M phosphate buffer, pH 6.8. (A) DPP curves of nonmodified DNA: 1, double-helical DNA; 2, thermally denatured DNA at the concentration of 0.030 mg/mL. (B and C) DPP curves of double-helical DNA modified by Ru-DHA (B) and Ru−CYM (C) at the following r_b values: 1, 0.0005; 2, 0.001; 3, 0.005; 4, 0.01. The arrows in parts A−C marked by II and III indicate the potentials *E* [against saturated calomel electrode (SCE)] at which native or denatured DNA samples yielded DPP peaks II or III, respectively (see the text). The procedure used to measure the height of the DPP peak (*x*) in the present work is shown in Figure 9A, curve 1. (D) Dependence of the relative height of the DPP peak II, ΔI, yielded by DNA modified by the ruthenium complexes on r_b: (■) Ru−BIP; (□) Ru−DHA; (○) Ru−CYM. The value of ΔI was calculated as the ratio of the peak height yielded by the modified DNA over the peak height yielded by the control (nonmodified) DNA.

the samples of DNA modified by Ru−BIP, Ru−DHA or Ru−THA at $r_b = 0.01$ to which 0.8% thermally denatured calf thymus DNA was added. For this sample, a small, more negative peak III on the DDP curve (recorded under conditions specified for curve 4 in Figure 9B) was clearly observed (not shown). Thus, the absence of peak III on the DPP curves of DNA modified by Ru−BIP, Ru−DHA, or Ru−THA suggests that these ruthenium complexes induce nondenaturational conformational distortions in DNA at relatively low levels of the global modification ($r_b \leq 0.01$), a behavior similar to that for the antitumor drug cisplatin and its antitumor analogues (*38*, *39*). In contrast, a relative increase in the intensity of peak II due to the global modification by Ru−CYM was also seen, but only at lower levels of the DNA modification ($r_b \leq 0.005$) (Figure 9C). At higher levels of DNA modification by Ru−CYM, a well-developed more negative DPP peak III, characteristic of the formation of single-stranded segments in double-helical

DNA, was observed. This finding supports the view that the DNA binding mode of Ru−CYM is different from that of the BIP, DHA, and THA ruthenium arene compounds and that the modification by Ru−CYM may even lead to denaturational distortions of DNA.

DISCUSSION

Our studies of the binding of the Ru(II) arene ethylenediamine complexes Ru−BIP, Ru−DHA and Ru−THA to natural double-helical CT DNA show that the reactions are about an order of magnitude faster than that for cisplatin (Figure 2). However, the rate of binding is markedly dependent on the type of arene ligand: binding of Ru−CYM, a complex with a single arene ring and methyl and bulky isopropyl substituents, binds to CT DNA more slowly than cisplatin. The binding is almost quantitative in each case (>90% after 24 h), and the DNA−Ru adducts are stable, with little loss of bound Ru after extensive dialysis. Transcription mapping experiments (Figure 3) have shown that guanine residues are the preferential binding sites when polymeric DNA is modified with Ru(II) arene complexes in a random fashion.

The selectivity of these pseudo-octahedral Ru(II) arene ethylenediamine complexes for G bases was also found previously in our studies of model reactions of monomeric nucleosides and nucleotides (*23*, *40*). Ru(II) in arene ethylenediamine complexes bind strongly to GN7, and C6O of G can then form a strong intramolecular H-bond with an NH of the coordinated ethylenediamine (*23*, *40*). For adenine and cytosine unfavorable interactions between the amino groups on these bases and the amino groups of coordinated en have been demonstrated (*23*, *40*). Binding to thymine N3 requires displacement of the N3H proton, which is not favorable at pH values in the physiological range and not accessible in double-helical DNA. Also, when the arene contains an extended π-electron system, as in biphenyl or the anthracene derivatives studied in this work, an additional stabilization of the interaction can arise from hydrophobic $\pi-\pi$ stacking of the coordinated arene ring system with the purine ring (*23*, *40*). Such hydrophobic $\pi-\pi$ stacking is enhanced when one or both partners are electron-poor (*41*). Ru(II) binding to G N7 decreases the electron density on the purine but increases electron density on Ru(II) which, in turn, enhances π-back-bonding of Ru(II) to the η^6-arene (*42*).

Slower reactions between CT DNA and Ru−CYM (Figure 1), or Ru−BEN, compared to Ru−BIP, Ru−DHA and Ru−THA are expected from our model studies of the kinetics of binding of these complexes to cyclic-3′,5′-GMP (*23*). Reactions of Ru−BIP, Ru−DHA and Ru−THA with cGMP are greater than three times faster. This difference may arise from the ability of the extended π system of the arene ligand in the faster-reacting complexes to take part in hydrophobic $\pi-\pi$ stacking interactions with the purine ring of G, as observed in model complexes. This is not possible for the monoarene complexes Ru−CYM or Ru−BEN, since the π-electron system of these ligands is fully involved in coordination to Ru(II). Hence, it is reasonable to suggest that hydrophobic interactions can contribute to the driving force for the binding of chloro Ru(II) arene complexes to double-helical DNA.

By analogy with the changes in the CD spectra and DPP behavior of DNA modified by cisplatin and antitumor-inactive transplatin or [Pt(dien)Cl]Cl, it is reasonable to suggest that the binding of Ru−BIP, Ru−DHA, and Ru−THA results in conformational alterations in double-helical DNA of nondenaturational character, as is the case for DNA modification by antitumor cisplatin (*38*). The CD and DPP results (Figures 4, 5, and 9) also suggest that the conformational changes induced in DNA by Ru−CYM and Ru−BEN are different from those induced by the other three ruthenium arene compounds and, in contrast, are of denaturational character, similar to the case of DNA modification by transplatin or monofunctional [Pt(dien)Cl]Cl complexes. The LD data (Figures S1 and S2) showed that Ru−CYM stiffened poly(dA-dT) DNA, while the other complexes bend it, and for Ru−BIP, Ru−DHA, and Ru−THA the main effect of binding to poly(dG-dC) and CT-DNA appears to be the induction of DNA bending, making it difficult to draw conclusions about the local orientations of the complexes on the DNA. The coordination of Ru−BIP, Ru−DHA, or Ru−THA to double-helical DNA also results in the appearance of a CD band centered around 370−380 nm (Figure 4A−C). The induced CD bands for achiral molecules bound to DNA arise either from coupling between the dipole transition moment of the nucleobases and the bound metal complex or from distortion of the geometry of the metal complex. Since binding of neither Ru−CYM (Figure 4D) nor Ru−BEN (Figure 5) to DNA gives rise to such a CD band in this region of the spectrum, and these complexes contain only a single ring, it seems likely that the metal-centered induced band is related either to intercalation of the extended arene ligands into DNA or to groove binding. The observation of such bands is well documented for other intercalating metal complexes (*43*, *44*). For example, the binding of [Pt(terpyridine)(CH_3)]$^+$ to DNA at $r_i = 0.1$ gives rise to positive CD bands at 315−340 nm, indicating possible intercalation of the terpyridine ligand (*45*). Similar CD bands were observed for binding of Ru(II) arene complexes to poly(dG-dC) as for CT DNA (Figure 5A,B), but not for poly(dA-dT) (Figure 5C). From model studies of mononucleotides (*23*, *40*), only weak binding to A or T is expected, accounting for the differences in binding to poly(dA-dT). From the assignments made by film LD experiments, the effects on the LD spectrum of poly(dA-dT) DNA are consistent with a binding mode involving insertion of the extended π-systems of Ru−DHA, Ru−THA, and Ru−BIP into the minor groove, giving a positive signal for the transitions in the 260−280 nm ligand y-polarized region and probably a negative signal for the 240−255 nm z-polarized region.

The CD changes observed for double-helical DNA modified by the Ru(II) arene compounds also correlate with the results of DNA unwinding (Figure 7 and Table 1) and competitive EtBr displacement (Figure 6) experiments. The monofunctional adducts of Ru−BIP, Ru−DHA, or Ru−THA are considerably more efficient in DNA−EtBr fluorescence quenching and in DNA unwinding than those of Ru−CYM. One plausible explanation for this observation may be an associated large additional contribution to fluorescence quenching and unwinding from intercalation of the extended arene ligand of the compounds Ru−BIP, Ru−DHA, or Ru−THA into the duplex or from other types of

noncovalent interaction of these complexes with DNA upon their monofunctional binding. The large unwinding angles of 15 or 19° produced by the platinum compounds *cis*-[Pt(NH_3)$_2$(N3-ethidium)Cl]$^{2+}$ and *cis*-[Pt(NH_3)$_2$(N8-ethidium)Cl]$^{2+}$, respectively, which incorporate the well-known DNA intercalator ethidium and which can form only monofunctional adducts with DNA, have been explained in this way (*28*). Thus, the arene moiety in the monofunctional adducts of the compounds Ru−BIP, Ru−DHA, and Ru−THA could be geometrically well positioned to intercalate between the base pairs of the helix, so producing also the induced CD bands (Figure 5). Consistent with this conclusion is the observation that the adducts of the compound Ru−CYM, which (like Ru−BEN) produces no induced CD bands in the visible spectrum and quenches DNA−EtBr fluorescence only slightly, unwinds DNA only by 7° (Figure 7D, Table 1), a similar behavior to that of the monofunctional adducts of [Pt(dien)Cl]Cl [unwinding angle 6° (*28*)]. Thus, the results of unwinding experiments support the view that the arene ligand in Ru−BIP, Ru−DHA, and Ru−THA interacts substantially with the double helix upon coordination of the ruthenium complex. Hence, these results strengthen the case for combined intercalative and monofunctional binding modes which may be facilitated by the fluxionality of the arene ligand in these complexes (*40*). On the other hand, it seems reasonable to suggest that the *p*-cymene and benzene ligands in the compounds Ru−CYM do not interact with the double helix in a similar way, thus also supporting a different DNA binding mode for these compounds in comparison with the other three Ru(II) arene complexes studied in the present work.

The arene−purine hydrophobic interactions and/or arene−base stacking involved in the binding of the arene compounds to double-helical DNA may also affect its melting behavior. Previously, three factors have been invoked to account for the thermal stability of DNA modified by ruthenium and platinum complexes: stabilizing effects of the positive charge on the ruthenium and platinum moieties and of DNA interstrand cross-links, and a destabilizing effect of conformational distortions such as intrastrand cross-links induced in DNA by ruthenium and platinum coordination. At least two additional factors may be involved in DNA stabilization by intercalators: (i) favorable stacking interactions between the base residues and the intercalator, and (ii) the separation of negative backbone charges inherent to intercalation (due to elongation and unwinding of DNA), that is, changes in solvent structure and the counterion distribution around the phosphate groups which may help to overcome electrostatics unfavorable for the hybridization of the strands of the duplex (*46*, *47*). The dependence of the transition melting temperature of DNA modified by nonintercalating platinum or ruthenium drugs on ionic strength can be explained by competing electrostatic effects as the salt concentration is varied (*48*). Under the incubation conditions, we expect all Ru(II) arene complexes to have produced monofunctional adducts so that the effect of interstrand cross-links need not be considered. Thus, the observed change in melting temperature will reflect the relative proportion and contribution of all limiting binding modes. Inherently, we predict that conformational distortions due to the formation of the adducts will destabilize the helix, as has been consistently observed in earlier studies with various ruthenium and platinum compounds.

At low ionic strength (0.01 M), it is reasonable to conclude that the increases in t_m due to the modification of DNA by Ru−BIP, Ru−DHA, or Ru−THA (Figure 8A−C) are caused by the positive charges on ruthenium {(arene)Ru(en)}$^+$ moieties and by the intercalation. An interesting and as yet unresolved question, therefore, is why the modification of DNA by these three Ru(II) arene compounds appears to result in smaller thermal stabilization or even destabilization if the melting curves are measured at high salt concentrations. It is possible that the smaller increase or decrease in t_m due to the modification by the compounds Ru−BIP, Ru−DHA, and Ru−THA and observed at high ionic strength is a consequence of conformational changes induced by the adducts of these Ru(II) arene compounds that then compensate more efficiently the "stabilizing" effects. At high salt concentration the stabilizing effects due to the modification of DNA by Ru−BIP, Ru−DHA, or Ru−THA are reduced, since the electrostatic effects of these compounds are apparently lowered with increasing concentration of Na^+ counterions.

As pointed out above, the melting behavior of DNA modified by the compound Ru−CYM is different (cf. parts A−C and D of Figure 8). Modification by Ru−CYM already decreases t_m at low concentrations of Na^+, indicating that the effects of the factors responsible for the thermal stabilization of DNA are noticeably reduced. As also mentioned above, two factors have been invoked to account for the increase of the thermal stability of DNA due to the modification by Ru(II) arene complexes: stabilizing effects of the positive charge on the ruthenium atom and those associated with intercalation of the arene ligand in these ruthenium compounds into the duplex. Hence, it seems reasonable to conclude that the effect of one or both these factors is markedly reduced so that the destabilization effect of conformational alterations induced by Ru−CYM predominates already at low salt concentrations. The results are consistent with the formation of a monofunctional adduct of Ru−CYM with DNA (coordination to G N7) and the absence of intercalation, as expected for this substituted single-ring arene ligand. Hence, the stabilization of DNA typical of intercalators (vide supra) would not be effective, for obvious reasons. In addition, the stabilizing effects of the positive charge on the ruthenium atom of the compound Ru−CYM might be considerably reduced due to a substantially different location of the ruthenium atom in the adduct of Ru−CYM relative to the DNA sugar−phosphate backbone. This location might be unfavorable from the viewpoint of the efficiency of the positive charge on the ruthenium atom to neutralize negative charges of DNA phosphate groups. The suggestion that the positive charge on the ruthenium atom in the adducts of Ru−CYM does not markedly contribute to the thermal stability of DNA is also consistent with the observation that the lowering of the melting temperature of DNA due to modification by Ru−CYM is almost independent of the Na^+ concentration (Figure 8D). Thus, the solution behavior of the DNA adducts of Ru(II) arene complexes appears interesting and merits further study.

Circular dichroism and polarographic analyses of DNA modified by Ru(II) arene anticancer compounds (Figures 4, 5, and 9) demonstrate that the formation of the monofunctional adducts of these complexes distorts the DNA confor-

mation. The compounds Ru−BIP, Ru−DHA, or Ru−THA, which contain arene ligands with extended π-systems, induce nondenaturational alterations in DNA. On the other hand, the adducts of Ru−CYM, which contains a single arene ring but with methyl and bulky isopropyl substituents, distort DNA more severely, so that even denaturational changes may occur. The mechanisms underlying the biological effects of antitumor metal-based drugs may also involve further "downstream" effects of damaged DNA. These effects include processing of DNA adducts of these compounds by cellular components, for example, recognition of the adducts by specific proteins, and their repair (*49*, *50*). This processing may also be affected by both the character of the conformational alterations induced in DNA and the resulting thermodynamic destabilization of this polynucleotide. For instance, minor 1,3-intrastrand cross-links of cisplatin, which probably do not contribute significantly to the antitumor effects of this drug, or the same adducts of its clinically ineffective trans isomer induce denaturational alterations in DNA (*51*). On the other hand, nondenaturational alterations occur in DNA as a consequence of the formation of the major 1,2-intrastrand cross-link of cisplatin (*52*). Interestingly, it has been shown that high-mobility-group (HMG)-domain proteins play an important role in the antitumor effect of cisplatin in several tumor cell lines (*49*, *50*), and while the 1,2-intrastrand cross-link of cisplatin is recognized by these proteins (*53*), the 1,3-intrastrand cross-links of cisplatin or transplatin are not (*50*, *53*). In addition, the 1,3-intrastrand cross-links of cisplatin are more readily removed from DNA than its major 1,2-intrastrand cross-links (*54*), and it has been shown (*49*, *55*) that enhanced repair of the adducts of cisplatin contributes to enhanced resistance to this drug. Thus, an intriguing eventuality for future research is to correlate different DNA binding modes of Ru−CYM and Ru−BEN and the other three Ru(II) arene compounds tested in the present work with their cytotoxicity in tumor cell lines.

In summary, the present work demonstrates that the concept based on Ru(II) arene complexes represents an interesting possibility for studies aimed at improving knowledge of the mechanisms underlying the biological effects of ruthenium compounds. Whether this concept is applicable to the development of new anticancer drugs has still to be examined.

SUPPORTING INFORMATION AVAILABLE

Flow linear dichroism spectra of CT, poly(dG-dC), and poly(dA-dT) DNA after reaction with Ru(II) arene complexes (Figure S1), and determination of transition polarizations of dihydroanthracene (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Johnson, N. P., Butour, J.-L., Villani, G., Wimmer, F. L., Defais, M., Pierson, V., and Brabec, V. (1989) *Prog. Clin. Biochem. Med. 10*, 1−24.
2. Brabec, V. (2000) in *Platinum-Based Drugs in Cancer Therapy* (Kelland, L. R., and Farrell, N. P., Eds.) pp 37−61, Humana Press Inc., Totowa, NJ.
3. Fichtinger-Schepman, A. M. J., Van der Veer, J. L., Den Hartog, J. H. J., Lohman, P. H. M., and Reedijk, J. (1985) *Biochemistry 24*, 707−713.
4. Eastman, A. (1987) *Pharmacol. Ther. 34*, 155−166.
5. Cleare, M. J. (1974) *Coord. Chem. Rev. 12*, 349−405.
6. Cleare, M. J. (1977) *J. Clin. Hematol. Oncol. 7*, 1−25.
7. Farrell, N. (1996) in *Metal Ions in Biological Systems* (Sigel, A., and Sigel, H., Eds.) pp 603−639, Marcel Dekker, Inc., New York, Basel, Hong Kong.
8. Brabec, V., Kasparkova, J., Vrana, O., Novakova, O., Cox, J. W., Qu, Y., and Farrell, N. (1999) *Biochemistry 38*, 6781−6790.
9. Farrell, N. (2000) in *Platinum-Based Drugs in Cancer Therapy* (Kelland, L. R., and Farrell, N. P., Eds.) pp 321−338, Humana Press Inc., Totowa, NJ.
10. Keppler, B. K. (1993) *Metal Complexes in Cancer Chemotherapy*, VCH Verlagsgesellschaft, VCH Publishers, Weinheim, New York.
11. Sava, G., and Bergamo, A. (2000) *Int. J. Oncol. 17*, 353−365.
12. Clarke, M. J. (2003) *Coord. Chem. Rev. 236*, 209−233.
13. Aird, R., Cummings, J., Ritchie, A., Muir, M., Morris, R., Chen, H., Sadler, P., and Jodrell, D. (2002) *Br. J. Cancer 86*, 1652−1657.
14. Morris, R. E., Aird, R. E., Murdoch, P. D., Chen, H. M., Cummings, J., Hughes, N. D., Parsons, S., Parkin, A., Boyd, G., Jodrell, D. I., and Sadler, P. J. (2001) *J. Med. Chem. 44*, 3616−3621.
15. Sava, G., Alessio, E., Bergano, A., and Mestroni, G. (1999) in *Topics in Biological Inorganic Chemistry. Metallopharmaceuticals I* (Clarke, M. J., and Sadler, P. J., Eds.) pp 143−169, Springer, Berlin.
16. Clarke, M. J., Zhu, F., and Frasca, D. R. (1999) *Chem. Rev. 99*, 2511−2533.
17. Sava, G., Pacor, S., Coluccia, M., Mariggio, M., Cocchietto, M., Alessio, E., and Mestroni, G. (1994) *Drug Invest. 8*, 150−161.
18. Novakova, O., Kasparkova, J., Vrana, O., van Vliet, P. M., Reedijk, J., and Brabec, V. (1995) *Biochemistry 34*, 12369−12378.
19. Barca, A., Pani, B., Tamaro, M., and Russo, E. (1999) *Mutation Res. 423*, 171−181.
20. Novakova, O., Hofr, C., and Brabec, V. (2000) *Biochem. Pharmacol. 60*, 1761−1771.
21. Malina, J., Novakova, O., Keppler, B. K., Alessio, E., and Brabec, V. (2001) *J. Biol. Inorg. Chem. 6*, 435−445.
22. Wang, F., Chen, H., Parsons, S., Oswald, I. D. H., Davidson, J. E., and Sadler, P. J. (2003) *Chem. Eur. J.*, in press.
23. Chen, H. M., Parkinson, J. A., Morris, R. E., and Sadler, P. J. (2003) *J. Am. Chem. Soc. 125*, 173−186.
24. Brabec, V., and Palecek, E. (1970) *Biophysik 6*, 290−300.
25. Brabec, V., and Palecek, E. (1976) *Biophys. Chem. 4*, 76−92.
26. Lemaire, M. A., Schwartz, A., Rahmouni, A. R., and Leng, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1982−1985.
27. Brabec, V., and Leng, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5345−5349.
28. Keck, M. V., and Lippard, S. J. (1992) *J. Am. Chem. Soc. 114*, 3386-3390.
29. Butour, J. L., and Macquet, J. P. (1977) *Eur. J. Biochem. 78*, 455−463.
30. Butour, J. L., Alvinerie, P., Souchard, J. P., Colson, P., Houssier, C., and Johnson, N. P. (1991) *Eur. J. Biochem. 202*, 975−980.
31. Bancroft, D. P., Lepre, C. A., and Lippard, S. J. (1990) *J. Am. Chem. Soc. 112*, 6860−6871.
32. Brabec, V., Boudny, V., and Balcarova, Z. (1994) *Biochemistry 33*, 1316−1322.
33. Rodger, A., and Norden, B. (1997) *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, New York, Tokyo.
34. Jenkins, T. C. (1997) in *Drug−DNA Interaction Protocols* (Fox, K. R., Ed.) pp 195−218, Humana Press Inc., Totowa, NJ.
35. Marini, V., Kasparkova, J., Novakova, O., Scolaro, L. M., Romeo, R., and Brabec, V. (2002) *J. Biol. Inorg. Chem. 7*, 725−734.
36. Palecek, E. (1983) in *Topics in Bioelectrochemistry and Bioenergetics* (Milazzo, G., Ed.) pp 65−155, John Wiley and Sons, Ltd., New York.
37. Palecek, E. (1971) *Methods Enzymol. 21*, 3−24.
38. Brabec, V., Kleinwächter, V., Butour, J. L., and Johnson, N. P. (1990) *Biophys. Chem. 35*, 129−141.
39. Vrana, O., Kleinwächter, V., and Brabec, V. (1984) *Experientia 40*, 446−451.
40. Chen, H. M., Parkinson, J. A., Parsons, S., Coxall, R. A., Gould, R. O., and Sadler, P. J. (2002) *J. Am. Chem. Soc. 124*, 3064−3082.
41. Janiak, C. (2000) *J. Chem. Soc., Dalton Trans.*, 3885−3896.
42. Stebler-Röthlisberger, M., Hummel, W., Pittet, P.-A., Bürgi, H.-B., Ludi, A., and Merbach, A. E. (1988) *Inorg. Chem. 27*, 1358−1363.

43. Long, E. C., and Barton, J. K. (1990) *Acc. Chem. Res. 23*, 271−273.
44. Lyng, R., Rodger, A., and Norden, B. (1991) *Biopolymers 31*, 1709−1720.
45. Arena, G., Monsu Scolaro, L., Pasternack, R. F., and Romeo, R. (1995) *Inorg. Chem. 34*, 2994−3002.
46. Maeda, Y., Nunomura, K., and Ohtsubo, E. (1990) *J. Mol. Biol. 215*, 321−329.
47. Bjorndal, M. T., and Fygenson, D. K. (2002) *Biopolymers 65*, 40−44.
48. Zaludova, R., Kleinwächter, V., and Brabec, V. (1996) *Biophys. Chem. 60*, 135−142.
49. Cohen, S. M., and Lippard, S. J. (2001) in *Prog. Nucleic Acid Res. Mol. Biol.* (Moldave, K., Ed.) pp 93−130, Academic Press Inc., San Diego, CA.
50. Brabec, V. (2002) in *Prog. Nucleic Acid Res. Mol. Biol.* (Moldave, K., Ed.) pp 1−68, Academic Press Inc., San Diego, CA.
51. Anin, M. F., and Leng, M. (1990) *Nucleic Acids Res. 18*, 4395−4400.
52. Gelasco, A., and Lippard, S. J. (1999) in *Topics in Biological Inorganic Chemistry. Metallopharmaceuticals I* (Clarke, M. J., and Sadler, P. J., Eds.) pp 1−43, Springer, Berlin.
53. Jamieson, E. R., and Lippard, S. J. (1999) *Chem. Rev. 99*, 2467−2498.
54. Zamble, D. B., Mu, D., Reardon, J. T., Sancar, A., and Lippard, S. J. (1996) *Biochemistry 35*, 10004−10013.
55. Brabec, V., and Kasparkova, J. (2002) *Drug Resist. Updates 5*, 147−161.

BI034933U

Spectroscopy 17 (2003) 653–661
IOS Press

Calibration and Standardisation of Synchrotron Radiation Circular Dichroism and Conventional Circular Dichroism Spectrophotometers

Andrew J. Miles [a], Frank Wien [a], Jonathan G. Lees [a], A. Rodger [b], Robert W. Janes [c] and B.A. Wallace [a,*]

[a] *Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK*
[b] *Department of Chemistry, Warwick University, Coventry CV4 7AL, UK*
[c] *School of Biological Sciences, Queen Mary, University of London, London E1 4NS, UK*

Abstract. Synchrotron radiation circular dichroism (SRCD) is an emerging technique in structural biology with particular value in protein secondary structure analyses since it permits the collection of data down to much lower wavelengths than conventional circular dichroism (cCD) instruments. Reference database spectra collected on different SRCD instruments in the future as well as current reference datasets derived from cCD spectra must be compatible. Therefore there is a need for standardization of calibration methods to ensure quality control. In this study, magnitude and optical rotation measurements on four cCD and three SRCD instruments were compared at 192.5, 219, 290 and 490 nm. At high wavelengths, all gave comparable results, however, at the lower wavelengths, some variations were observable. The consequences of these differences on the spectrum, and the calculated secondary structure, of a representative protein (myoglobin) are demonstrated. A method is proposed for standardising spectra obtained on any CD instrument, conventional or synchrotron-based, with respect to existing and future databases.

1. Introduction

Synchrotron radiation circular dichroism (SRCD) spectroscopy is a method that was first developed more than 20 years ago [1,2] but which is just beginning to realise its full potential in the study of protein secondary structures. It permits the collection of much lower wavelength data than conventional circular dichroism (cCD) spectroscopy due to the high light flux available from the synchrotron source. Indeed, recently a protein spectrum in aqueous solution was obtained with a lower wavelength limit of $\leqslant$154 nm [3], as compared to a practical limit of $\sim$185–190 nm on conventional instruments. It has been suggested that the extra information from the vacuum ultraviolet (VUV) data will allow a greater discernment of structural features, including information on folds and motifs [4]. Creation of a new reference database of protein CD spectra, including the lower wavelength data which will enable fold recognition studies, is in progress [3].

There are currently a number of operational SRCD instruments world-wide [5], and several more will be coming on-line in the near future [6]. Therefore establishment of a consistent set of calibration standards and protocols for SRCD and cCD is now essential. Instrumental calibrations of magnitude,

*Corresponding author. E-mail: ubcg25a@mail.cryst.bbk.ac.uk.

polarisation, and wavelength are needed, as are precise measurements of optical cell pathlengths and protein concentrations, in order to obtain the correct measurements of ellipticities that are necessary for secondary structural analyses [7–9]. Since the currently available reference databases for protein secondary structural analyses were derived from measurements made on cCD instruments [10], and future reference databases will be measured on SRCD instruments but must also be usable with cCD data, it is especially important that SRCD and cCD instruments are calibrated to the same standards [6]. Establishing protocols for calibration will also be important in the context of "good practice" procedures for protein drug standardisation and characterisation in the pharmaceutical industry where CD is currently being employed to confirm batch-to-batch reproducibility of protein structure.

Conventional circular dichroism instruments are routinely calibrated for amplitude at 290 nm using camphor sulphonic acid (CSA) [11] or ammonium camphorsulphonate (ACS) [12]. However, calibrations should be done at more than one wavelength [13] in order to cover the wavelength range measured in a protein spectrum, and to demonstrate the linearity of the response. Other candidates that have been proposed for calibration standards are the lower wavelength (192.5 nm) CSA peak [11], pantolactone (PL) [14] with a peak at 219 nm, and cobalt (III) *tris*-ethylenediamine (Co-en), with a maximum at 490 nm [15], which also provides a good standard for the visible region.

As a preliminary attempt at standardisation, this study compares optical rotation magnitude measurements for a number of compounds on three SRCD instruments with those obtained on four cCD instruments. In addition, spectra of the protein myoglobin were also measured on all these instruments in order to examine the effects of instrument variation on the spectrum of a standard protein. The calculated secondary structures derived from these spectra were then compared in order to determine the consequences of the variations. Finally, a method is proposed for standardisation of protein spectra obtained on any instrument, which will "correct" the spectra and enable empirical analyses based on existing and future database spectra, using the values obtained for four standards on these instruments.

2. Materials and methods

2.1. Materials

(D)-(−)-pantolactone, of 99.99% purity (Sigma-Aldrich Company Ltd), was dissolved at a concentration of 0.0115 M in 18.2 MΩ deionized water (dH_2O). (+)-Camphor-10-sulphonic acid, 99% purity (Sigma-Aldrich Company Ltd), was dissolved in dH_2O at a concentration of 0.047 M, determined by the absorbance at 285 nm assuming an extinction coefficient of 34.5 M^{-1} cm^{-1} [16]. Co[(ethylenediamine)$_3$]Cl_3·NaCl·$6H_2O$, the kind gift of Peter W. Thulstrup (from Roskilde University, Denmark), was dissolved in dH_2O at a concentration of 33.4 mM, as determined by the absorbance at 464 nm, using an extinction coefficient of 84 M^{-1} cm^{-1} [15]. The CSA sample was diluted to concentrations ranging from 0.9× to 0.1× to assay the CD signal as a function of concentration, as a test of the linearity of the response. The samples of CSA, PL and Co-en were stored in the dark at 4°C and used within 2 weeks of preparation. In the first set of experiments, the same samples (using the same CD cells) were tested on most of the machines (4 cCDs and 1 SRCD) within that 2-week period. These samples were retested on the first machine at the end of this time to ensure that no changes had taken place. Because of constraints associated with beamtime scheduling which did not permit measurements on the other two SRCDs to be done during the initial two week period, another set of samples was prepared in the same way and was used for measurements on them, and for comparison, on one of the cCDs that had been used for the first set of measurements.

Horse skeletal myoglobin (ICN Biochemicals) was allowed to dissolve in dH_2O overnight at a concentration of $\sim$8 mg/ml. One ml of the solution was then dialysed against 50 ml of dH_2O for 2 hrs to reduce the salt content, degassed and centrifuged at 5000 $\times$ g to separate any undissolved material. The final concentration of the protein was determined to be 7.43 mg/ml by replicate quantitative amino acid analyses.

2.2. CD and SRCD spectroscopy

The following parameters were set on all instruments: bandwidth 1 nm, step size 0.2 nm, temperature 25°C. Spectra were collected on the Aviv 62ds and 215 instruments and on all the SRCD instruments in step scan mode with an averaging (dwell) time of 1 second. At least 3 measurements were made and averaged for each sample on each instrument. The Jasco J-600 and J-715 instruments were run in continuous scan mode, averaging over 8 accumulations with a response time of 0.5 s. The PL, CSA, and Co-en spectra were all collected using a 1 mm pathlength Suprasil cell (Hellma UK Ltd). A dH_2O baseline, collected using the same cell and the same parameters, was subtracted from all spectra. All myoglobin spectra were obtained using the same cylindrical 0.001 cm pathlength demountable cell (Hellma UK Ltd) held in a specially adapted cell holder; baselines of dialysate were collected in the same cell.

SRCD measurements were made at the SRS Daresbury (UK) on beamline CD12; at ISA (Denmark) on beamline UV1; and at BESSY2 (Germany) on beamline 3m_NIM1_C. Where possible, the parameters were set to the values used on the cCD instruments. 3m_NIM1_C as yet had no bandwidth control, so the slits were fully open, and due to limited availability of beamtime, only 2 repeats were done. Because two of the beamlines had upper wavelength limits of $<$400 nm, measurements on the Co-en sample were not possible on these instruments. The SRCD instrument on beamline CD12 records intensities as arbitrary counts. Therefore, the literature value for CSA at 290 nm was assumed and the other values calculated from this for the CD12 data.

The pathlengths of the optical cells used were determined using the absorbance of a 1 mM solution of potassium chromate ($\varepsilon_{372} = 4830\ M^{-1}\,cm^{-1}$ (Chemical Rubber Company)) and by the interference fringe method [17] for the (nominally) 0.1 and 0.001 cm cells, respectively.

2.3. Secondary structure calculations

The secondary structural analyses used the DICHROWEB (http://www.cryst.bbk.ac.uk/cdweb) interactive webserver [8] with the CONTIN [10,18] algorithm and database 6 [8]. The normalised root mean square deviation (NRMSD) parameter [19] was calculated as a means of assessing the goodness-of-fit.

2.4. Standardisation method

For each of the instruments for which measurements were available for all four standards, a standardisation procedure was applied to the myoglobin spectrum collected on that instrument. This was done as follows: for each standard, the literature delta epsilon value was divided by the delta epsilon value measured on that instrument. Then a second-order polynomial was fit to the ratios at the four different wavelengths, and the ratio values at all other wavelengths (R_λ) were interpolated from the ratio versus wavelength curve. Then for each myoglobin data point, the scaled delta epsilon value $\Delta\varepsilon_{s\lambda} = R_\lambda \cdot (\Delta\varepsilon_\lambda)$ was plotted as the "standardised" spectrum. The procedure is detailed in the Annex of this paper.

3. Results

One advantage of this study is that the comparative measurements could be done on equivalent samples (close in time and prior to degradation) using the same CD cell (and thus the same pathlength), by the same investigator, thereby eliminating many possible sources of systematic errors, and enabling a true comparison of the instrument characteristics.

3.1. Accuracy of optical rotation magnitude measurements

Pantolactone has a $\Delta\varepsilon$ at 219 nm of -4.9 mdeg M^{-1} cm^{-1} [20] therefore a signal of 186 mdeg was expected at this wavelength for the 0.0115 M solution used in this work.

CSA has a $\Delta\varepsilon$ at 290 nm of 2.37 mdeg M^{-1} cm^{-1} [20] therefore a signal of 367 mdeg was expected for the 0.047 M solution used in this work. $\Delta\varepsilon$ for the second CSA peak at 192.5 nm is 4.72 mdeg M^{-1} cm^{-1} [11]. Therefore a peak of 146 mdeg was expected from a 5 times dilution of the CSA sample. It should be noted that the CSA (and all peaks) varied in wavelength position between the different instruments: The low wavelength CSA peak was found at positions ranging from 190.5 to 194.0 nm, and the high wavelength peak was between 289.1 and 293.2 nm. Thus calibration of wavelength position is another important parameter to consider for instrument standardisation (Miles & Wallace, in prep.). In this paper, however, the ellipticity values reported are always for the peak maximum, regardless of what the peak wavelength position was.

Assuming enantiomeric purity, with a $\Delta\varepsilon$ at 490 nm of 1.89 mdeg M^{-1} cm^{-1} [15], the Co-en sample used in this work should give rise to a CD signal of 208 mdeg. The results obtained on all instruments are displayed graphically in Fig. 1a. The "literature value" curve is established by a simple fit (Excel) to the four data points and is for visualisation purposes only. For clarity, all the cCD measurements are shown as open symbols, and all the SRCD measurements are shown as filled symbols. A given symbol signifies data measured on a single instrument.

Since instruments are generally routinely calibrated using the CSA 290 nm peak, it is not surprising that they all correspond well to each other and to the literature value at this wavelength. However, the deviations are considerably larger at lower wavelengths.

At 490 nm, all ellipticity measurements were within 6% of the literature value. Variations for single machines across all four calibration standards range from 0 to 3% for one cCD instrument to 4 to 18% for another cCD instrument, with the SRCD instruments mostly falling in the middle range (0 to 10%). The variations include values that are both too large and too small, with no particular pattern being discernable. Also, there is no clear trend of deviation between cCDs or the SRCDs. The most significant result is that none of the instruments vary by a constant amount from the literature values across all wavelengths, so no simple scale factor can be applied. This is in complete contrast to the common practice whereby an instrument is calibrated at a single point and it is assumed the whole wavelength range is thereby calibrated.

At all three calibration wavelengths in the UV, two of the cCDs are calibrated to within 3% of the expected value. This is close to the 2% error level estimated from repeated measurements. One cCD displays significant deviation at 219 nm (-10%) and another deviates by $+11\%$ at 192.5 nm and $+18\%$ at 219 nm. One SRCD instrument was found to have a deviation of $\sim$8% at all the wavelengths measured, whilst another was within 4% of the literature values at three wavelengths, but deviates at 192.5 nm by $>10\%$, hence its 192.5/290 peak ratio is the highest value measured, 2.20.

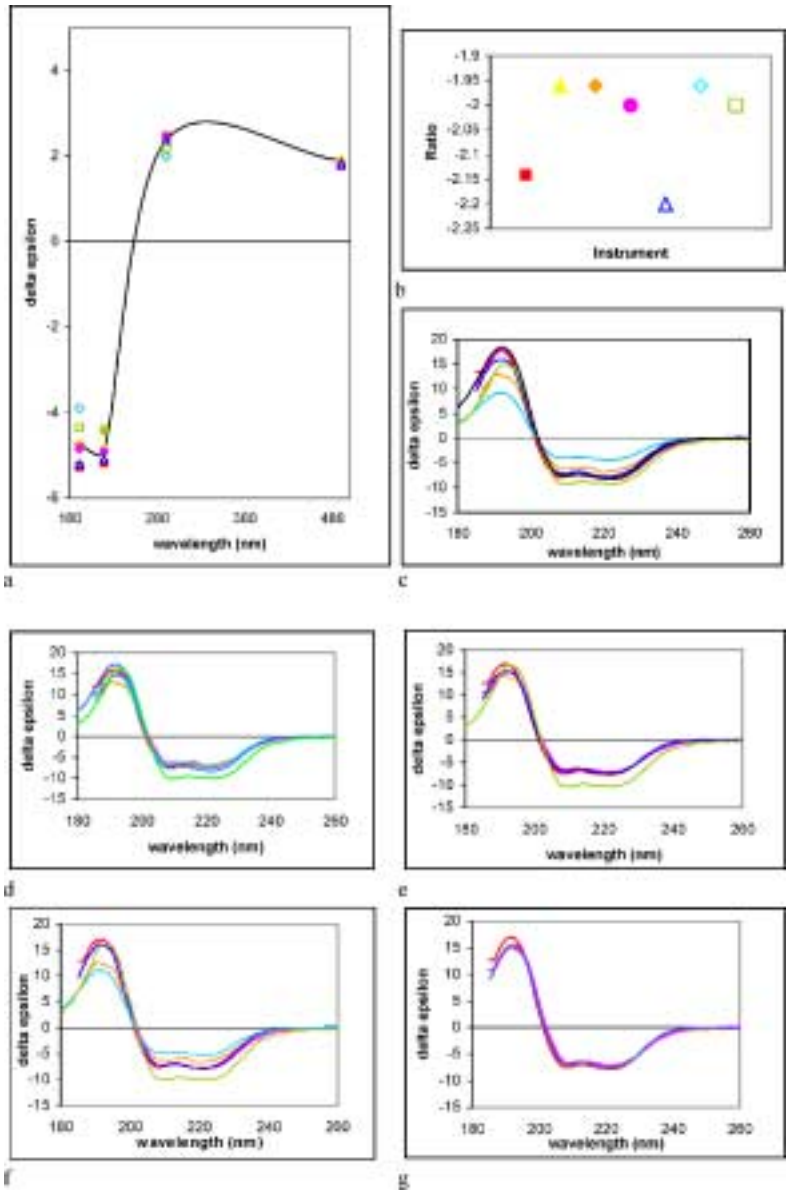


Fig. 1. Variations between different instruments: (a) Amplitude variations of four cCD and three SRCD instruments for three standard compounds (CSA, PL, and Co-en) at four wavelengths. The "literature value" curve is established by a simple fit to the values, and is for visualisation purposes only. cCD measurements are shown as open symbols, SRCD measurements are shown as filled symbols. A given symbol signifies data measured on a single instrument. (b) Plot of ($\Delta\varepsilon_{192.5\ \mathrm{nm}}/\Delta\varepsilon_{290\ \mathrm{nm}}$) ratios of CSA for the cCD and SRCD instruments, using the same symbols as in Fig. 1a. (c) Uncorrected myoglobin spectra obtained on cCD and SRCD instruments. All spectra were obtained using the same concentration of protein and the same cell (and hence the same pathlength). The spectra from the cCD instruments are shown in red/orange/yellow/purple; the spectra from the SRCD instruments are in light and dark blue and green (in each case, the same colour as used for the symbols in Figs 1a,b). The spectrum of myoglobin from the reference database [10] is shown in black, for comparison. (d–f) Myoglobin spectra from various instruments scaled to calibration points at (c) 290 nm, (d) 219 nm, and (e) 192.5 nm, respectively. It can be seen that there is still considerable variation between spectra obtained on different instruments, with no clear segregation into cCD and SRCD populations. (g) Convergence of the myoglobin spectra after standardisation by the multiwavelength ("four-point") method proposed in this paper.

3.2. Linearity of detector response

It is important to demonstrate that the instruments produce linear responses across the range of ellipticity values measured, in order for the CSA dilution calculations to be correct. This was assayed by measuring a series of dilutions of the CSA sample, and plotting ellipticity versus concentration (data not shown). All instruments surveyed exhibited linear responses up to ellipticity values of 300 mdeg.

Another important issue to consider is the total absorbance (often measured as high tension (HT)) of the sample. When the sample absorbance is too high, the number of photons reaching the detector becomes too small for accurate measurement and the CD signals cannot be trusted. The cut-off must therefore be determined on each individual instrument. In this study, all measurements were done under conditions where the absorbance was still within the accurately measurable range.

3.3. CSA signal ratios

The ratio of the ellipticities of the two CSA peaks ($\Delta\varepsilon_{192.5\ \mathrm{nm}}/\Delta\varepsilon_{290\ \mathrm{nm}}$) is often used as a simple "two-point" means of calibrating the instrument, as opposed to the absolute measurement of the ellipticity of a single peak. Reported values in the literature range from $\sim$1.90 to 2.20 (i.e., [11,17,21–23]). In this study, the values for the ratio ranged from 1.96 to 2.20 (Fig. 1b). However, this parameter on its own is not particularly useful, as the absolute magnitude may be incorrect, whilst the ratio is reasonable. Also, there is no clear standard for the "correct" value, with significant variations possible due to, amongst other sources, CSA purity, light-induced degradation, and instrument stray light. Hence, we consider the magnitudes of the individual peaks may be more suitable measures of the CD calibration. However, they are dependent on having accurate values for the extinction coefficients and ellipticity values for the standard materials.

3.4. Effects of instrumental variation on protein spectra

The variations between instruments for the standards are considerable in the critical far UV spectral range used for protein analyses. The obvious consequences of this is that protein spectra measured on the various instruments will differ considerably both in magnitude and shape (the latter because the deviations from true signals are different for different wavelengths) from each other (Fig. 1c) and from the whale myoglobin spectrum in the reference database (black curve, Fig. 1c). Therefore the simple scaling of all the spectra by a factor determined at one wavelength will not produce equivalent spectra. To illustrate this, all the spectra were normalised to the literature values at each of the single UV wavelengths (192.5, 219 and 290 nm), and replotted (Figs 1d–f). The most successful of these was the scaling to the 219 peak, but in this case the spectra showed some variation, especially the one outlier (green) that had very different values for the standards at all wavelengths. Nevertheless, none of the single wavelength scalings produced very good correspondences. An alternative, the four-point standardisation method, described above, resulted in very close correspondences with each other (Fig. 1g) and with the spectrum in the reference database.

3.5. Effects of instrumental variation on calculated secondary structures

Two instruments (one cCD and one SRCD) produce spectra nearly identical to the reference database spectrum of myoglobin, so it is no surprise that they produce similar calculated secondary structures.

However, the unscaled spectra (Fig. 1c) for the other instruments produce dramatically different calculated secondary structures (ranging from 0.44 to 0.83, average 0.71), where the value calculated by the DSSP algorithm from the crystal structure was 0.75 [24]. This result demonstrates that without proper cross-calibration, empirical methods for secondary structural analyses will not work for either cCD or SRCD measurements. However, after the four-point standardisation scaling, the resulting spectra produced a narrower range of secondary structures (range from 0.70 to 0.74, average = 0.72), close to the crystal structure value. On this basis, we propose that the scaling method described may be a useful one for both cCD and SRCD data.

4. Discussion

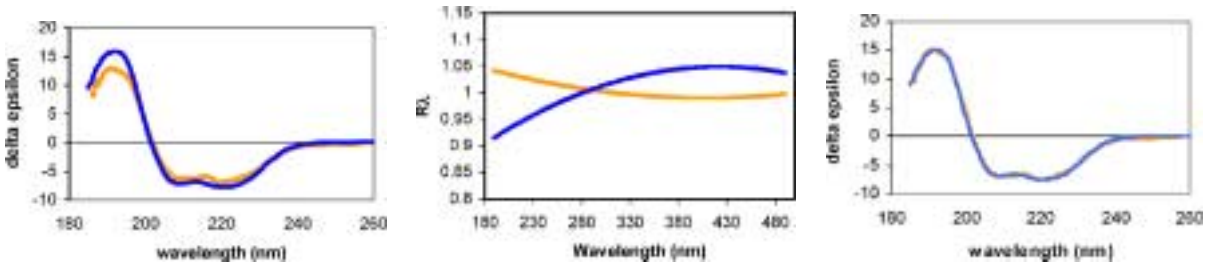
4.1. Comparisons of cCD and SRCD spectra

It can be seen that there is considerable variation between spectra obtained on various instruments, with no clear segregation into cCD and SRCD populations, and no clear trend between or within the types of instruments. There is very close correspondence between one of the cCD instruments (red) and one of the SRCD instruments (light blue). In a previous study, it was suggested that at the very low wavelength limits, conventional spectra differ slightly from SRCD spectra [25]; this was attributed to the necessity of fully opening the slits in the cCD to achieve high enough light flux, which ultimately allowed through light of other wavelengths, and seriously questions the use of data at wavelengths $\leqslant$185 nm from cCD instruments. The purpose of the present study, however, was to examine the correlations at wavelengths where both instruments are operating optimally, i.e. above $\sim$190 nm.

4.2. Proposed standardisation method for calibration corrections

The variations between the measurements of the standards on the various instruments are reflected in the spectra of the protein sample. The overall magnitudes and the relative peak magnitudes are significantly different in spectra from the different instruments (Fig. 1c), and most experimental spectra are at variance with the corresponding spectrum from the reference database (in black, Fig. 1c). These spectral differences result in large differences and inaccuracies in the calculated secondary structures derived from them. The variations could also lead to problems in use of CD for "good practice" quality control analyses of pharmaceutical proteins if uncorrected data from different instruments were compared. As a result, it is clear that there needs to be some procedure that can be used to standardise/calibrate the spectra taken on one instrument with the spectra taken on other instruments and in the reference databases.

The established method for calibration of cCD instruments uses only the value for the CSA calibration at 290 nm to normalise the spectra on a given instrument. That this is actually not a very suitable method (see Fig. 1f) is unsurprising given that this peak is outside the far UV wavelength range used for measuring protein spectra. Normalisations using either the 192.5 nm CSA or the 219 nm PL peaks are somewhat better (Figs 1d and 1e) but still do not take into account the wavelength-dependent variations. Hence we propose a new procedure which utilises the ratios of the measured and expected delta epsilon values at each wavelength to develop a model function for corrections to be applied to spectra. This has proved to be a much more successful method of calibration than any single wavelength approach, resulting in superposable spectra from the various instruments and with the corresponding spectrum from the reference dataset (Fig. 1g).



(a) (b) (c)

Fig. 2. Examples of application of the standardisation method to data collected on two different CD instruments (blue and orange curves): (a) uncorrected spectra of myoglobin obtained on the two CD instruments; (b) plots of R_λ vs. wavelength for the two instruments, derived from fits to the four standard measurements; (c) corrected (according to the method described in the Annex) spectra of myoglobin obtained on the two CD instruments.

Acknowledgements

This work was supported by grants B13586 and B14225 from the BBSRC to BAW. JL was supported by a BBSRC CASE studentship with GlaxoSmithKline, and AM was supported by an MRC studentship. Beamtime access at ISA was enabled by an EC Access to Research Infrastructure Action of the Improving Human Potential Programme to Soren Pape Moller (Aarhus University). We thank the following for technical assistance at the SRCD beamlines: Dr. David Clarke (CD12 beamline, SRS Daresbury), Dr. Soren Vronning Hoffmann (UV1 beamline, ISA), and Dr. Peter Baumgartel (beamline 3m_NIM1_C, BESSY2). We thank Dr. Peter Sharratt of the PNAC facility, University of Cambridge, for the quantitative amino acid analyses.

Annex: Protocol for the four-point standardisation method

1. Measure the CD spectrum of the protein of interest (Fig. 2a).
2. Measure the CD spectra for the three standards (CSA, PL, Co-en) at the following four wavelengths: 192.5, 219, 290, 490 nm.
3. For the above measurements on the standards, calculate the ratio (R_λ) of the literature value for the ellipticity divided by the measured value for the ellipticity.
4. Fit a second order polynomial to the four R_λ vs. wavelength data points (Fig. 2b).
5. Determine the value of the polynomial (R_λ) for each wavelength over the whole spectral range.
6. Multiply the CD spectrum of the protein of interest by the R_λ value at each wavelength and plot the resulting standardised spectrum (Fig. 2c).

References

[1] J.C. Sutherland, E.J. Desmond and P.Z. Takacs, Versatile spectrometer for experiments using synchrotron radiation at wavelengths greater than 100 nm, *Nucl. Instr. Meth.* **172** (1980), 195–199.
[2] P.A. Snyder and E.M. Rowe, The first use of synchrotron radiation for vacuum ultraviolet circular dichroism measurements, *Nucl. Instr. Meth.* **172** (1980), 345–349.
[3] B.A. Wallace, F. Wien, A.J. Miles, J.G. Lees, P. Evans, S.V. Hoffman, G.J. Wistow and C. Slingsby, Biomedical applications of synchrotron radiation circular dichroism spectroscopy: identification of mutant proteins associated with disease and development of a reference database for fold motifs, *J. Chem. Soc.* (2003), in press.

[4] B.A. Wallace and R.W. Janes, Synchrotron radiation circular dichroism spectroscopy of proteins: secondary structure, fold recognition and structural genomics, *Curr. Opin. Chem. Biol.* **5** (2001), 567–571.
[5] B.A. Wallace, Synchrotron radiation circular dichroism spectroscopy as a tool for investigating protein structures, *J. Synch. Rad.* **7** (2000), 289–295.
[6] B.A. Wallace, First International Workshop on SRCD Spectroscopy, *Synchrotron Radiation News* **15** (2002), 20–22.
[7] B.A. Wallace and C.L. Teeters, Differential absorption flattening optical effects are significant in the circular dichroism spectra of large membrane fragments, *Biochemistry* **26** (1987), 65–70.
[8] A. Lobley, L. Whitmore and B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics* **18** (2002), 211–212.
[9] L. Whitmore, A. Miles and B.A. Wallace, in preparation.
[10] N. Sreerama and R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set, *Anal. Biochemistry* **282** (2000), 252–260.
[11] G.C. Chen and J.T. Yang, Two-point calibration of circular dichrometer with d-10-camphorsulfonic acid, *Anal. Lett.* **10** (1977), 1195–1207.
[12] T. Takakuwa, T. Konno and H. Meguro, A new standard substance for calibration of circular dichroism: ammonium d-10-camphorsulfonate, *Anal. Sci.* **1** (1985), 215–218.
[13] K. Tuzimura, T. Konno, H. Meguro, M. Hatano, T. Murakimi, K. Kashiwabara, K. Saito, Y. Kondo and T.M. Suzuki, A critical study of the measurement and calibration of circular dichroism, *Anal. Biochem.* **81** (1977), 167–174.
[14] T. Konno, H. Meguro and K. Tuzimura, D-Pantolactone as a circular dichroism (CD) calibration, *Anal. Biochem.* **67** (1975), 226–232.
[15] A.J. McCaffery and S.F. Mason, The electronic spectra, optical rotatory power and absolute configuration of metal complexes: the dextro-*tris* (ethylenediamine)cobalt (III) ion, *Mol. Phys.* **6** (1963), 359–371.
[16] T.M. Lowry, *Optical Rotatory Power*, Dover Publications, New York, 1964, p. 407.
[17] J.P. Hennessey, Jr. and W.C. Johnson, Jr., Experimental errors and their effect on analysing circular dichroism spectra of proteins, *Anal. Biochem.* **125** (1982), 177–188.
[18] I.H.M. Van Stokkum, H.J.W. Spoelder, M. Bloemendal, R. Van Grondelle and F.C.A. Groen, Estimation of protein secondary structure and error analysis from CD spectra, *Anal. Biochemistry* **191** (1990), 110–118.
[19] D. Mao, E. Wachter and B.A.Wallace, Folding of the mitochondrial H^+-ATPase proteolipid in phospholipid vesicles, *Biochemistry* **21** (1982), 4960–4968.
[20] P.H. Schippers and H.P.J.M. Dekkers, Direct determination of absolute circular dichroism data and calibration of commercial instruments, *Anal. Chem.* **53** (1981), 778–782.
[21] W.C. Johnson, Jr., Circular dichroism instrumentation, in: *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York, 1996, pp. 635–652.
[22] S.Y. Venyaminov and J.T. Yang, Determination of protein secondary structure, in: *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, New York, 1996, pp. 69–107.
[23] W.C. Johnson, Jr. and I. Tinoco, Circular dichroism of polypeptide solutions. I the vacuum ultraviolet, *J. Amer. Chem. Soc.* **94** (1972), 4389–4390.
[24] B.A. Wallace, J. Lees, A.J.W. Orry, A. Lobley and R.W. Janes, Analyses of circular dichroism spectra of membrane proteins, *Protein Sci.* **12** (2003), 875–884.
[25] J.G. Lees and B.A. Wallace, Synchrotron radiation circular dichroism and conventional circular dichroism spectroscopy: a comparison, *Spectroscopy* **16** (2002), 121–125.

Eur. J. Biochem. **270**, 3345–3352 (2003) © FEBS 2003 doi:10.1046/j.1432-1033.2003.03710.x

An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on the hydrophobicity of the environment

Miguel San Miguel[1], Rachel Marrington[2], P. Mark Rodger[2], Alison Rodger[2] and Colin Robinson[1,*]

[1]*Department of Biological Sciences and* [2]*Department of Chemistry, University of Warwick, Coventry, UK*

The Tat system catalyzes the transport of folded globular proteins across the bacterial plasma membrane and the chloroplast thylakoid. It recognizes cleavable signal peptides containing a critical twin-arginine motif but little is known of the overall structure of these peptides. In this report, we have analyzed the secondary structure of the SufI signal peptide, together with those of two nonfunctional variants in which the region around the twin-arginine, RRQFI, is replaced by KKQFI or RRQAA. Circular dichroism studies show that the SufI peptide exists as an unstructured peptide in aqueous solvent with essentially no stable secondary structure. In membrane-mimetic environments such as SDS micelles or water/trifluoroethanol, however, the peptide adopts a structure containing up to about 40% α-helical content. Secondary structure predictions and molecular modelling programs strongly suggest that the helical region begins at, or close to, the twin-arginine motif. Studies on the thermal stability of the helix demonstrate a sharp transition between the unstructured and helical states, suggesting that the peptide exists in one of two distinct states. The two non-functional peptides exhibit almost identical spectra and properties to the wild-type SufI peptide, indicating that it is the arginine sidechains, and not their contribution to the helical structure, that are critical in this class of peptide.

Keywords: signal peptide; twin-arginine translocation; Tat system; protein transport; SufI.

The twin-arginine translocation (Tat) system operates in the cytoplasmic membranes of most free-living bacteria and in the thylakoid membranes of plant chloroplasts [1–3]. Operating alongside Sec-type translocases, the Tat system functions in the transport of proteins bearing cleavable N-terminal signal peptides (RR-signal peptides) in which a twin-arginine motif plays a central role [4,5]. The substrate proteins are recognized by a membrane-bound translocase and subsequently transported, at least in some cases, in a fully folded state [6,7]. The prime role of the Tat system appears to be in the transport of proteins, which are either obliged to fold prior to translocation, or which fold too tightly for the Sec system to accommodate. Examples of the former category include periplasmic proteins that are exported only after binding any of a range of complex redox cofactors, such as molybdopterin or FeS centres [8–11]. These cofactors are inserted in the cytoplasm by complex enzymatic processes, and it has been argued that this must necessitate the export of these proteins in a largely, if not fully, folded form.

Substrates bearing RR-signal peptides are recognized by a membrane-bound translocase that consists minimally of TatABC in most cases. Critical genes encoding these subunits have been identified in bacteria, particularly *Escherichia coli* [10–13], and in plants [14–16], and a TatABC complex has been purified from detergent-solubilized *E. coli* membranes [17]. The size of the purified complex has been estimated to be in the order of 500–600 kDa, suggesting the presence of numerous copies of each subunit [17]. The TatBC subunits form a tight core subcomplex in a strict 1 : 1 ratio [17] and studies on the thylakoid system indicate that this subcomplex forms the initial binding site for substrates, with TatA recruited at a later stage [18,19]. Consistent with this scenario, important residues in *E. coli* TatC have been identified on the cytoplasmic side of the membrane [20,21].

Although the importance of the twin-arginine motif has been firmly established by mutagenesis studies, RR-signal peptides have not been characterized in structural respects. The peptides are probably too small to form folded globular domains but with other types of targeting signal it has been shown that functionality is strictly dependent on the formation of specific secondary structures. Prominent examples include Sec-type signal peptides and the presequences of imported mitochondrial proteins (e.g [22,23]). In this report, we have analyzed a typical *E. coli* RR-signal peptide and we show that the structure of the peptide differs dramatically according to environment. Implications for the translocation process are discussed.

Correspondence to C. Robinson, Plant Biochemistry Laboratory, Department of Plant Biology, Royal Veterinary and Agricultural University, 40, Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark.
Fax: + 44 2476523701, Tel.: + 44 2476523557,
E-mail: CRobinson@bio.warwick.ac.uk

Abbreviations: MD, molecular dynamics; TFE, trifluoroethanol.

**Present address*: Plant Biochemistry Laboratory, Department of Plant Biology, Royal Veterinary and Agricultural University, 40, Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark.

(Received 15 March 2003, revised 2 June 2003, accepted 9 June 2003)

Experimental procedures

Materials

Peptides, purified by HPLC, were purchased from Alta Bioscience (Birmingham, UK).

Circular dichroism

Circular dichroism (CD) spectra are the simplest indication of protein and peptide secondary structure. CD spectra were collected using a Jasco J-715 spectropolarimeter equipped with a single sample Peltier thermostatting unit. Spectra were averaged over eight scans collected with 1 nm data intervals, 1 nm bandwidths, 0.5 s response time and 200 nm·s^{-1} scan speed. The CD spectra as a function of temperature were collected by monitoring at a single wavelength (222 nm, 2 nm bandwidth) with an 8-s response time and a ramp rate of 1 °C·min^{-1}. At each of 20, 30, 40, 50, 60, 70, 80, and 90 °C, the temperature ramp was held while a single wavelength scan with 4-s response time was collected. All samples were made up to a concentration of 0.10 mg·mL^{-1} by weighing a peptide sample on a seven-figure balance and adding the appropriate volume of solvent. In the case of the mixed organic–aqueous solvents, the solvent was fully mixed prior to adding to the dry peptide. Organic solvent volumes were measured using a syringe. Aqueous solution volumes were measured using micropipettes. α-Helical content was estimated using a $\Delta\varepsilon$ value of ≈ 12 mol^{-1}·dm^{-3}·cm^{-1} for a 100% helical peptide [24] and also by applying the protein CD structure-fitting program, CDSSTR [25]. The latter approach may only be used as a guide as its database does not properly account for random coil structures.

Modelling

Energy minimization calculations were performed using QUANTA/CHARMM version 28 (Accelerys Inc., Cambridge UK) and molecular dynamics simulations with the DL_POLY package [26]. Coordinates and force field were constructed within QUANTA/CHARMM and then exported and converted to DL_POLY format using in-house software. Test calculations were performed on representative single configurations using both CHARMM and DL_POLY to check that no differences were observed between any of the energy calculations. Initial configurations for the peptide were constructed in two different ways: (a) using the PHD structure prediction method [27], and (b) from the lowest energy configuration obtained from minimizing configurations generated during a vacuum molecular dynamics (MD) simulation at 500 K. The PHD method gives an α-helix covering 16 amino acids, while the MD method gives a random coil structure (see Figs 4 and 5). Initial structures were optimized using the conjugate gradient method (CHARMM). The peptide was then inserted into a solvent box (65 Å and previously equilibrated at 300 K, 1 atm), all solvent molecules that overlapped with the peptide removed, and three additional solvent molecules converted to Cl^- ions to compensate for the +3 charge of SUFI; this resulted in 3173 water molecules, or 990 trifluoroethanol (TFE) molecules, in a periodic truncated octahedral simulation box. The system was then relaxed by (a) performing a 5-ps MD simulation at 300 K, 1 atm in which the peptide was treated as a rigid body, and then (ii) performing a 2-ps MD simulation with a fully flexible peptide at 2 K; these stages served to remove any strain introduced on solvation without destroying the initial secondary structure. A further 6 ns simulation was then accumulated to study the response of the secondary structure. Secondary structure was analyzed using the STRIDE program [28]. All MD simulations were performed at constant temperature and pressure (NPT) using the Nosé–Hoover method with thermostat and barostat relaxation constants of 0.5 and 1.0, respectively, and a timestep of 2 fs. The peptide and TFE were modelled with the CHARMM potential, and water with the SPC model. We note that the published TFE potentials do have some inadequacies [29], and that the CHARMM potential can underestimate the stability of α-helices relative to other force fields [30], but overall the model is reasonable. Long range forces were truncated at 10 Å, and the reaction field method used to correct for long-range electrostatic effects.

Secondary structure prediction

Secondary structure was predicted using the PSIPRED [31], JPRED [32], PROF [33] and PHD [27] secondary structure prediction programs.

Results

Structures of wild-type and nonfunctional SufI signal peptides

The overall aim in this study was to analyze the structural characteristics of a wild-type RR-signal peptide in both aqueous solvent and membrane-mimicking environments, as well as in some intermediate environments. The objective was to identify the secondary structure characteristics immediately after synthesis (in aqueous medium, as the targeting process is post-translational at least in the vast majority of cases) and once bound either to the membrane or the translocase (see below). A possibility addressed in this study was that the twin-arginine motif somehow may contribute significantly to the secondary structure, and we therefore also analyzed mutant variants that are not recognized by the Tat system. In this way, we sought to determine whether the significance of the twin-arginine motif stems from the nature of the sidechains and/or its ability to promote a given form of secondary structure.

We chose SufI as an example of a typical *E. coli* Tat substrate. SufI is synthesized with an apparently typical RR-signal peptide with the sequence shown in Fig. 1. Two variants are also shown; in one (SufI-KK), the twin-arginine motif is replaced by twin-lysine, which causes either a complete or near-complete block in translocation by the Tat pathway in both chloroplasts and bacteria [4,5,34]. In the other, the region around twin-arginine motif, RRQFI, is replaced by RRQAA (SufI-AA). While the effects of this double mutation have not been tested in bacterial systems, studies on the thylakoid system have demonstrated the critical importance of a highly hydrophobic residue at the second or third positions after the twin-arginine motif [35].

Fig. 1. Structure of the SufI signal peptide. Shown are the amino acid sequences of the *E. coli* SufI twin-arginine signal peptide, together with the sequences of two mutant versions used in this study. The essential twin-arginine motif is shown as bold italics.

SufI	MSLS***RR***QFIQASGIALCAGAVPLKASA	(mw= 2746 Da)
SufI-KK	MSLSKKQFIQASGIALCAGAVPLKASA	(mw = 2690 Da)
SufI-AA	MSLS***RR***QAAQASGIALCAGAVPLKASA	(mw = 2628 Da)

Bacterial RR-signal peptides invariably contain such a residue, typically phenylalanine, leucine, isoleucine or valine, at one or both of these positions [8] and the SufI-AA peptide shown in Fig. 1 is therefore strongly predicted to be nonfunctional. The SufI peptide has been synthesized chemically and shown to competitively inhibit the Tat-dependent transport of proteins into inverted *E. coli* vesicles [36], and to interact with Tat complexes in detergent [37]. This indicates that the peptide is functional in isolation and we used the same peptide in the present study in order to examine the structure of the signal at various stages. The KK mutant peptide does not affect binding of wild-type substrate suggesting that the RR motif is required for substrate binding.

The SufI RR-signal peptide is unstructured in aqueous solvent but α-helical in membrane-mimetic environments

The secondary structure of the SufI RR signal and the mutant variants was analyzed by CD. Figure 2 shows their structures in aqueous solvent (pH = 7) and two membrane-mimetic environments, namely SDS micelles and water/TFE systems. A range of alcohols in which the peptide was soluble were also investigated and spectra in methanol and ethanol shown for SufI. The data obtained with the wild-type peptide in water (Fig. 2A) show a spectrum characteristic of an unstructured peptide with a negative maximum below 200 nm and no signal at 220 nm. The spectra in the other solvent systems are typical of α-helices, though of

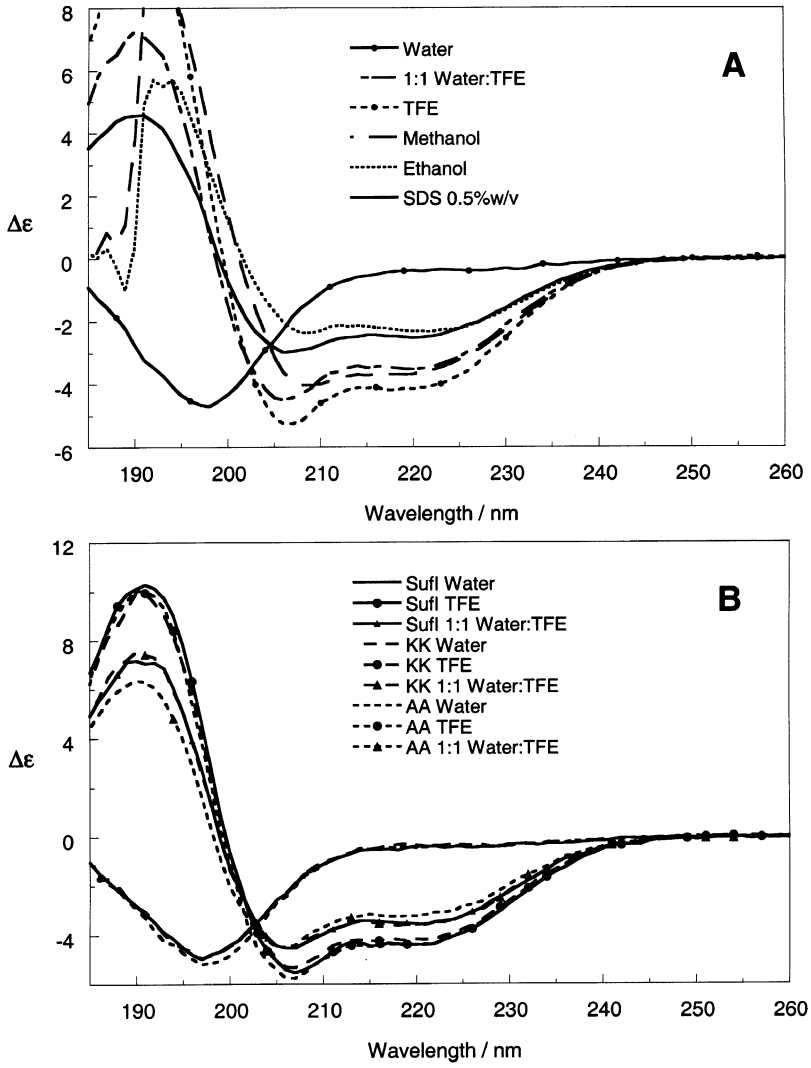


Fig. 2. CD spectra of SufI signal peptides in aqueous and apolar solvents. (A) CD spectra were taken using the wild-type SufI signal peptide in solvents as indicated in the figure. Further details are given in Experimental procedures. (B) Comparison of the wild-type peptide with the SufI-KK and SufI-AA peptides (denoted as 'KK' or 'AA', respectively, in inset to graph). CD spectra were taken in water, water/TFE (50 : 50 v/v) and TFE.

magnitude significantly less than expected for 100% helix ($\Delta\varepsilon \approx 12$ mol^{-1}·dm^{-3}·cm^{-1} at 208 nm and 222 nm) even in TFE where the helical content is ≈45%. The alcohols ethanol, propan-1-ol and butan-1-ol gave the same spectrum within experimental error down to 195 nm, with an α-helix content of ≈25%. Methanol induces a noticeably more helical content (≈ 40%), similar in magnitude to a 50% TFE/water solvent. A 0.5% w/v SDS solution resulted in a structure similar to that observed with ethanol.

Figure 2B shows the spectra obtained with the SufI-KK and SufI-AA peptides. In both cases, the spectra are almost identical to those of the wild-type indicating that the twin-arginine motif does not contribute specific α-helical properties in this context. Compared with SufI, the 208 nm region of the spectra indicates a slight increase in helical character for SufI-AA and a slight decrease for SufI-KK in TFE, whereas at 222 nm in water/TFE SufI-AA has less intensity and SufI-KK is the same as SufI. This suggests that the mutant helices are perhaps of slightly different form. The SufI-AA random coil spectrum (water) also differs slightly from the others.

Thermal stability of the α-helix: evidence for two distinct states

The CD spectra of SufI collected in 50 : 50 water/TFE as a function of temperature show a well-defined nonzero isosbestic point at 201 nm (Fig. 3A). This indicates that the system is in one of two states, presumably α-helical and random coil. The fact that there is no sharp transition during the temperature ramp (Fig. 3B) suggests that we have a temperature dependent equilibrium between the two states rather than any sort of concerted transition. The situation with the two mutants is very similar, though there is only an approximate isosbestic point in each case, with KK being slightly the worse (data not shown). Although there are not enough data to enable us to determine the cause of this, the most likely cause is a slight variation in helical forms present in the solution.

Location of the helical segment(s) in hydrophobic milieu

The SufI RR-peptide clearly contains substantial amounts of helical structure in apolar environments and, because the CD data do not indicate the location of this structure within the 27-residue peptide, we used additional methods in order to pinpoint the likely location(s). First, the PSIPRED, JPRED, PROF AND PHD secondary structure prediction programs were used, and typical predictions are shown in Fig. 4A. All four programs predict substantial α-helical content (26, 48, 56 and 59%, respectively) and it is notable that these regions encompass the twin-arginine motif in each case, with the RR motif usually positioned at or near the beginning of the helical section.

As a complementary technique we carried out a detailed molecular modelling simulation of SufI RR in both water and TFE. Simulations were performed using two different initial peptide conformations, one taken from the PHD structure prediction (59% α-helix; diagram 1 of Fig. 4B) and one with essentially a random coil secondary structure (the lowest energy structure from the conformational search; not shown). The simulated secondary structure from these two initial configurations was found to have converged after about 1 ns, indicating that the subsequent structures were stable and not merely an artefact of the simulation timescale, nor of the starting geometry.

The results of the simulations are in semiquantitative agreement with the CD experiments. No α-helix was found to persist in water, and indeed the 16 amino acid helix in the PHD initial structure completely disappeared within ≈ 200 ps. Typical peptide conformations from the end of a simulation in water are shown in diagrams 3 and 4 of Fig. 4B. In contrast, the long time behaviour of the simulations in TFE showed about 20% α-helix (not shown). This value is lower than the CD experiments suggest (≈ 45%). Given the reservations about the TFE potential discussed earlier, this probably indicates that the TFE potential is slightly too hydrophilic in character so that only the most significant α-helix-forming tendency emerges in the simulation. While some fluctuation in the length of the helix was observed, the RR motif was consistently found at the beginning of the helix.

Discussion

Numerous RR-signal peptides have been characterized in terms of primary sequence and it is now clear that, in most cases, they conform to a standard basic model in which three distinct domains can be identified: a charged N-terminal (N-) domain, hydrophobic core (H-) domain and more polar C-terminal domain terminating with a consensus motif that specifies cleavage by the processing peptidase. Key structural determinants are located at the boundary between the N-domain and the hydrophobic core. In plants, RR-signal peptides contain an essential twin-arginine motif plus a highly hydrophobic residue two or three residues towards the C-terminus. Typically, the sequence RRXXL/F/I/M or RRXL/F/I/M is found [35]. In *E. coli*, a slightly different consensus is observed, namely RRxFLK [8]. However, *E. coli* RR-signal peptides are efficiently recognized by plant thyalkoids indicating a highly conserved interaction between the signal peptide and the Tat translocation apparatus [38,39].

A surprising point is that Tat-dependent signal peptides closely resemble Sec-type signals in overall terms. Sec signals likewise contain the three domains described above and the only immediately notable difference is that the basic residue(s) at the N–H domain junction can be either arginine or lysine. In bacteria, Sec signals also tend to be more hydrophobic than RR-signal peptides. Nevertheless, these similarities raise the possibility that the signals may not always direct targeting by the correct translocation pathway, at least initially.

While the primary sequences of Tat-type signal peptides are now relatively well established, very little is known of their secondary structures. This is an important issue for two reasons. Firstly, it is now clear that the twin-arginine motif plays a crucial role in the translocation process and this could be either (a) because its sidechains are specifically recognized by the Tat translocase, and/or (b) because it confers a specific secondary structure determinant that is important in the context of the remainder of the peptide. Second, the targeting of Tat substrates is in most (if not all) cases an obligatorily post-translational

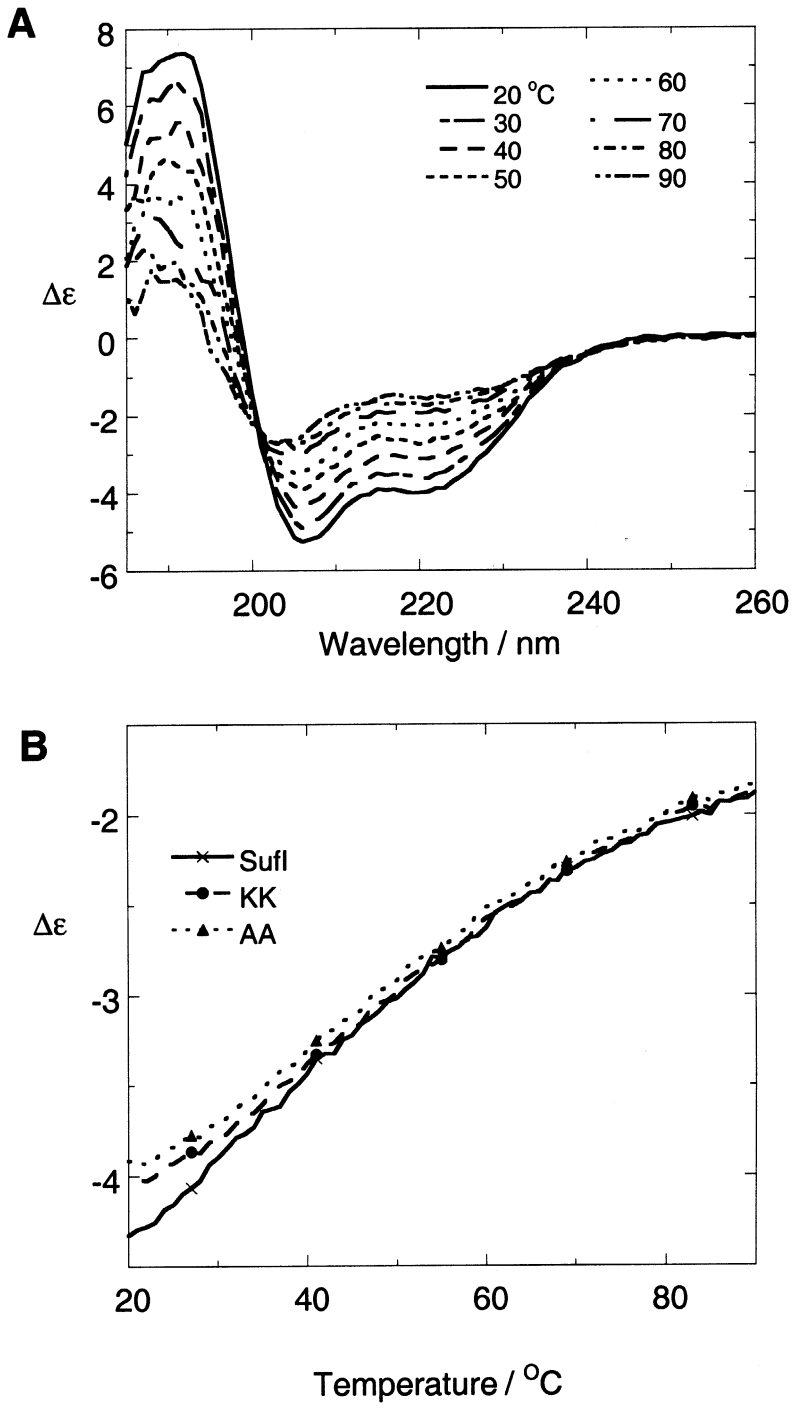


Fig. 3. Thermal stability of the helical regions in SufI signal peptides. (A) The wild-type SufI peptide in TFE/water (50 : 50 v/v) was analyzed by CD and spectra were taken at the temperatures indicated. (B) Similar CD spectra were recorded using the SufI-AA and SufI-KK peptides (not shown) and the graph shows plots of the change in Δε at 201 nm as a function of temperature for all three peptides.

process, raising the possibility that the signal may adopt different structures before and after reaching the target membrane.

In this study, we have used the SufI signal peptide as a model and the data show that this signal can adopt two radically different structures. In aqueous solution, our data show that the signal contains essentially no stable secondary structure. CD is particularly sensitive for the detection of α-helices, and the data show quite clearly that the signal has no stable helical elements, presumably because of competition for hydrogen bonding by water molecules. After release from the ribosome, Tat signal peptides are therefore likely to exist as unstructured peptides. In more hydrophobic environments, namely 50 : 50 water/TFE or SDS micelles, the structure is very different and the wild-type SufI peptide contains approximately 40 and 25% α-helical structure, respectively. Both secondary structure predictions and MD simulations of SufI in TFE (explicit solvent) strongly suggest that the helical structure is located in the centre section of the peptide, starting either just before or at

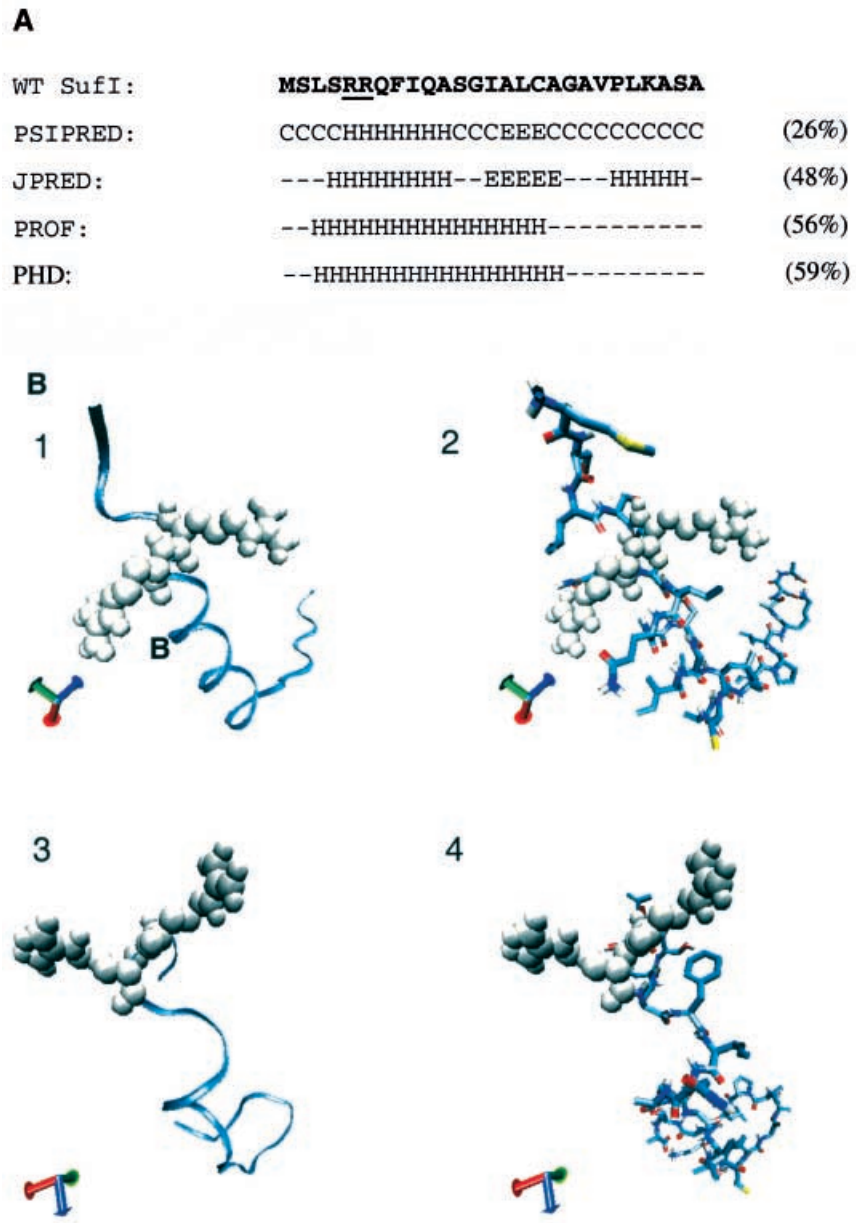


Fig. 4. Predicted secondary structure in the SufI signal peptide. (A) Secondary structure for the wild-type SufI peptide (WT) predicted using the PSIPRED, JPRED, PROF and PHD programs. C indicates random coil, H an α-helix and E an extended or sheet motif. The percentage helical content of the peptides is shown on the right. (B) 3D representations (ribbons mode on the left and bonds mode on the right) of the initial structures for the WT peptide obtained from the PHD program (1 and 2) and final configurations from an MD simulation in water (3 and 4); the twin-arginine motif has been depicted using a space-filling representation.

the twin-arginine motif. The MD simulations also indicate that the formation of an α-helix is very rapid in a hydrophobic environment.

While the twin-arginine is strongly predicted to lie at or near the starting point for the helical domain, our data indicate that it does not itself dictate the formation or extent of the helical sequence in a specific manner because a mutant version of the peptide containing twin-lysine contains almost exactly the same amount of α-helical structure. This finding strongly suggests that the real significance of the twin-arginine motif is concerned with protein–protein interactions involving the two extensive arginyl sidechains.

On the basis of these data we suggest the following model for the structures of Tat signal peptides during the overall translocation process (Fig. 5). Immediately after synthesis, the signal peptide probably exists in equilibrium between unstructured peptide and a more structured peptide containing substantial α-helix in the core region. While in aqueous solvent, the unstructured peptide predominates, to the point where the helical form is essentially undetectable. There is no evidence for the involvement of any soluble

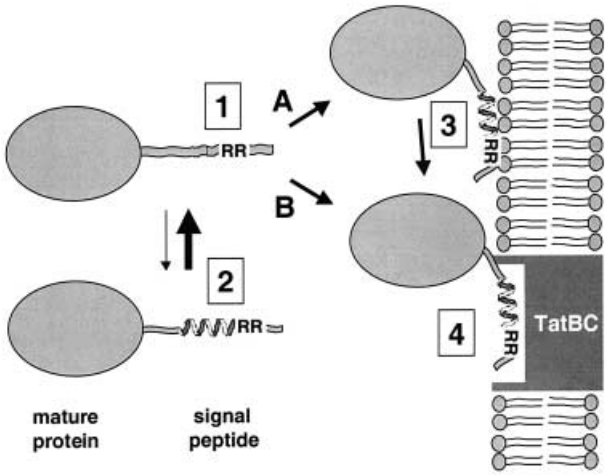


Fig. 5. Models for the structures of signal peptides during the overall Tat-dependent translocation process. Immediately after synthesis, the signal peptide is exposed to an aqueous environment where it is in equilibrium between the unstructured (1) or substantially α-helical (2) conformations. The equilibrium is strongly in favour of the unstructured peptide, probably because of competition for H-bonding by water molecules. Thereafter, targeting to the Tat translocase can occur by one of two routes. In pathway A, the peptide interacts with the membrane and α-helix formation is promoted within the interfacial region, after which the precursor protein migrates laterally until contact with the translocase is made. In B, the precursor protein interacts directly with the translocase, in which α-helix formation may be promoted within the binding site.

chaperone-type proteins and we therefore believe that the next protein–protein interaction is with the Tat translocase, probably by one of two pathways as shown in Fig. 5. In pathway A, the helical structure is promoted by interaction with the target membrane. There is now good evidence that α-helix formation is strongly promoted within the membrane interfacial region [40] and the precursor protein may then migrate laterally on the membrane surface until it contacts the translocase. In pathway B, the signal interacts directly with the translocase and the helical conformation is generated either during entry into the interfacial region or after the initial interaction with the binding site. Any form of typical binding site/groove probably also favours the formation of secondary structure, and the enormously specific signal peptide–translocase interaction would be difficult to achieve if the peptide were largely unstructured. We therefore contend that the signal peptide must be in the helical form when docked onto the translocase binding site. A precedent is the interaction of mitochondrial targeting peptides with the Tom20 receptor, where the targeting signal is likewise unstructured in aqueous solution but in the form of an amphipathic α-helix when bound to the receptor [41,42]. Further work should determine whether the same applies to the Tat-dependent translocation process.

Acknowledgements

This work was supported by Engineering and Physical Sciences Research Council grant GR/R36503 to CR and PMR.

References

1. Robinson, C. & Bolhuis, A. (2001) Protein targeting by the twin-arginine translocation pathway. *Nat. Rev. Mol. Cell. Biol.* **2**, 350–355.
2. Dalbey, R.E. & Kuhn, A. (2000) Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Annu. Rev. Cell Dev. Biol.* **16**, 51–87.
3. Yen, M.R., Tseng, Y.H., Nguyen, E.H., Wu, L.F. & Saier, M.H. Jr (2002) Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch. Microbiol.* **177**, 441–450.
4. Chaddock, A.M., Mant, A., Karnauchov, I., Brink, S., Herrmann, R.G., Klösgen, R.B. & andRobinson, C. (1995) A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the ΔpH-dependent thylakoidal protein translocase. *EMBO J.* **14**, 2715–2722.
5. Stanley, N.R., Palmer, T. & Berks, B.C. (2000) The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J. Biol. Chem.* **275**, 11591–11596.
6. Clark, S.A. & Theg, S.M. (1997) A folded protein can be transported across the chloroplast envelope and thylakoid membranes. *Mol. Biol. Cell* **8**, 923–934.
7. Hynds, P.J., Robinson, D. & Robinson, C. (1998) The Sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J. Biol. Chem.* **273**, 34868–34874.
8. Berks, B.C. (1996) A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* **22**, 393–404.
9. Santini, C.-L., Ize, B., Chanal, A., Müller, M., Giordano, G. & Wu, L.-F. (1998) A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J.* **17**, 101–112.
10. Weiner, J.H., Bilous, P.T., Shaw, G.M., Lubitz, S.P., Frost, L., Thomas, G.H., Cole, J.A. & Turner, R.J. (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**, 93–101.
11. Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C. & andPalmer, T. (1998) Overlapping functions of components of a bacterial Sec-independent export pathway. *EMBO J.* **17**, 3640–3650.
12. Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C. & Palmer, T. (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273**, 18003–18006.
13. Sargent, F., Stanley, N.R., Berks, B.C. & Palmer, T. (1999) Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J. Biol. Chem.* **274**, 36073–36082.
14. Voelker, R. & Barkan, A. (1995) Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J.* **14**, 3905–3914.
15. Settles, M.A., Yonetani, A., Baron, A., Bush, D.R., Cline, K. & Martienssen, R. (1997) Sec-independent protein translocation by the maize Hcf106 protein. *Science* **278**, 1467–1470.
16. Walker, M.B., Roy, L.M., Coleman, E., Voelker, R. & Barkan, A. (1999) The maize tha4 gene functions in sec-independent protein transport in chloroplasts and is related to hcf106, tatA, and tatB. *J. Cell. Biol.* **147**, 267–276.
17. Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C. & Robinson, C. (2001) TatB and TatC form a structural and functional unit of the twin-arginine translocase of *Escherichia coli*. *J. Biol. Chem.* **276**, 20213–20219.
18. Cline, K. & Mori, H. (2001) Thylakoid ΔpH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J. Cell Biol.* **154**, 719–729.

19. Mori, H. & Cline, K. (2002) A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid Delta pH/Tat translocase. *J. Cell Biol.* **157**, 205–210.
20. Allen, S.C.H., Barrett, C.M.L., Ray, N. & Robinson, C. (2002) Essential cytoplasmic domains in the *Escherichia coli* TatC protein. *J. Biol. Chem.* **277**, 10362–10366.
21. Buchanan, G., Leeuw, E., Stanley, N.R., Wexler, M., Berks, B.C., Sargent, F. & Palmer, T. (2002) Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol. Microbiol.* **43**, 1457–1470.
22. Roise, D., Theiler, F., Horvath, S.J., Tomich, J.M., Richards, J.H., Allison, D.S. & Schatz, G. (1988) Amphiphilicity is essential for mitochondrial presequence function. *EMBO J.* **7**, 649–653.
23. Jones, J.D., McKnight, C.J. & Gierasch, L.M. (1990) Biophysical studies of signal peptides: implications for signal sequence functions and the involvement of lipid in protein export. *J. Bioenerg. Biomembr.* **22**, 213–232.
24. Woody, R.W. (1994) *Circular Dichroism: Principles and Applications* (Nakanishi, K., Berova, N. & Woody, R.W., eds). VCH, New York.
25. Johnson, W.C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins Struct. Funct. Genet.* **35**, 307–312.
26. Smith, W., Yong, C.W. & Rodger, P.M. (2002) DL_POLY: application to molecular simulation. *Mol. Sim.* **28**, 385–471.
27. Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.* **266**, 525–539.
28. Frishman, D. & Argos, P. (1995) Knowledge-based protein secondary structure assignment. *Proteins* **23**, 566–579.
29. Chitra, R. & Smith, P.E.J. (2001) A comparison of the properties of 2,2,2-trifluoroethanol and 2,2,2-trifluoroethanol/water mixtures using different force fields. *J. Chem. Phys.* **115**, 5521–5530.
30. Rodziewicz-Motowidio, S., Brzozowski, K., Legowska, A., Liwo, A., Silbering, J., Smoluch, M. & Rolka, K.J. (2002) Conformational solution studies of neuropeptide gamma using CD and NMR spectroscopy. *J. Pept. Sci.* **8**, 211–226.
31. McGuffin, L.J., Bryson, K. & Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405.
32. Cuff, J.A. & Barton, G.J. (1999) Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins* **40**, 502–511.
33. Ouali, M. & King, R.D. (2000) Cascaded multiple classifiers for secondary structure prediction. *Prot. Sci.* **9**, 1162–1176.
34. Henry, R., Carrigan, M., McCaffrey, M., Ma, X. & Cline, K. (1997) Targeting determinants and proposed evolutionary basis for the Sec and the delta pH protein transport systems in chloroplast thylakoid membranes. *J. Cell Biol.* **136**, 823–832.
35. Brink, S., Bogsch, E.G., Edwards, W.R., Hynds, P.J. & Robinson, C. (1998) Targeting of thylakoid proteins by the ΔpH-driven twin-arginine translocation pathway requires a specific signal in the hydrophobic domain in conjunction with the twin-arginine motif. *FEBS Letts* **434**, 425–430.
36. Alami, M., Trescher, D., Wu, L.F. & Muller, M. (2002) Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in *Escherichia coli*. *J. Biol. Chem.* **277**, 20499–20503.
37. de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S., Muller, M., Sargent, F., Palmer, T. & Berks, B. (2002) Oligomeric properties and signal peptide binding by *Escherichia coli* Tat protein transport complexes. *J. Mol. Biol.* **322**, 1135–1146.
38. Mori, H. & Cline, K. (1998) A signal peptide that directs non-Sec transport in bacteria also directs efficient and exclusive transport on the thylakoid Delta pH pathway. *J. Biol. Chem.* **273**, 11405–11408.
39. Wexler, M., Bogsch, E.G., Palmer, T., Robinson, C. & Berks, B.C. (1998) Targeting signals for a bacterial Sec-independent export system direct plant thylakoid import by the delta pH pathway. *FEBS Letts.* **431**, 339–342.
40. White, S.H. & Wimley, W.C. (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struc.* **28**, 319–365.
41. Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T. & Kohda, D. (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* **100**, 551–560.
42. Epand, R.M., Hui, S.W., Argan, C., Gillespie, L.L. & Shore, G.C. (1986) Structural analysis and amphiphilic properties of a chemically synthesized mitochondrial signal peptide. *J. Biol. Chem.* **261**, 10017–10020.

Publication 87

Groovy DNA recognition: new agents for gene regulation?

Mike Hannon and Alison Rodger

Department of chemistry, University of Warwick, Coventry CV4 7AL, UK

Tel: +44 (0)24 7652-4107/3243. E-mail: m.j.hannon@warwick.ac.uk; a.rodger@warwick.ac.uk

In nature, many biomolecules bind to DNA sequences by inserting themselves into the major groove of the double helix – a region largely ignored by chemists when developing small molecule recognition agents. The authors have used simple metal-ligand interactions to assemble synthetic supramolecular cylinders which bind to this major groove and wrap up DNA into tight packages. Possible applications of such an intra-molecular coiling process might include preventing DNA processing, enabling rogue genes to be suppressed

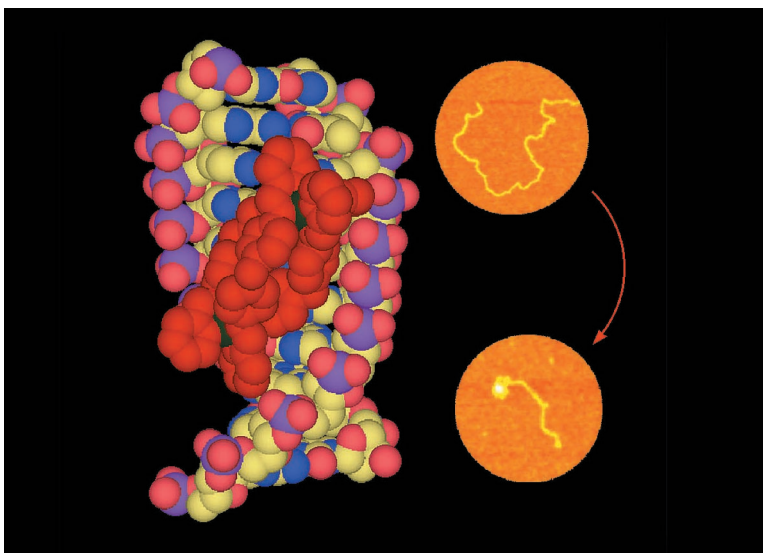


Figure 1. A graphical represenatation showing the cylinder binding to the major groove and the dramatic coiling it induces

The genome project has not only caught the imagination of the press and public, but in a very real sense is driving the direction of current and future biological research with new emerging scientific disciplines, such as genomics and proteomics. Genome (and proteome) information has enormous implications and potential for developing medical treatments that it is hoped will revolutionise the quality (and possibly length) of life for humans in the 21st century. Much current drug research focuses on agents that target the action of proteins through which disease is manifested. However, the post-genomic environment presents the opportunity to control disease at source by stimulating or preventing the processing of specific genes on our DNA. This is particularly important for combating a range of genetic-based diseases, such as cancers, viruses and Down's Syndrome or the other trisomies. To achieve such ambitious goals, an arsenal of different types of molecular tools will be required, which recognise the genetic code in a sequence-selective fashion.

NATURAL DNA RECOGNITION

Within biological systems, sequence-specific code recognition is generally achieved by the surface motifs of proteins. Their interactions are frequently non-covalent in character (hydrogen bonds, Van der Waals, hydrophobic and electrostatic interactions) and often via the DNA major groove, which is particularly attractive for DNA recognition since:

- the size and shape of the major groove of B-DNA[1] varies with base sequence (more so than the minor groove); and
- a number of additional hydrogen-bond donor and acceptor units are positioned in the major groove, and their relative spatial dispositions are dependent on the DNA base sequence.

Examples of such protein DNA-recognition include transcriptional regulators, which often involve cylindrical binding units, such as alpha helices or zinc fingers, that insert into the major groove and may kink the DNA. Other biomolecules that can selectively recognise DNA include oligonucleotides (synthetic or natural short lengths of DNA) which can bind in the major groove forming DNA triplexes. Oligonucleotide analogues (such as the neutral peptide nucleic acids) can also act in a similar fashion, although more commonly they act through strand displacement.

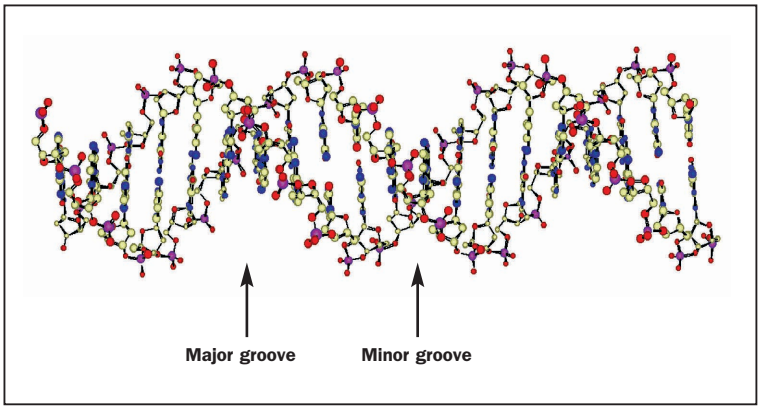


Figure 2. Diagrammatic representation of B-DNA illustrating its major and minor grooves

By way of contrast, synthetic small molecule recognition agents designed by chemists tend to interact with DNA in quite different ways and the attractive major groove favoured by biomacromolecules is rarely a target. Synthetic molecules that achieve sequence selectivity include well-known minor groove binders such as amide-linked imidazole/pyrole oligomers and polymers developed particularly by Dervan and co-workers.[2] These molecules have been shown to compete with some regulatory binding proteins and to repress gene expression. Alternatively many small molecules containing extended planar aromatic ring systems act via intercalation where they insert between the DNA base pairs (the gap between the pairs opens up to allow the insertion and this lengthens the DNA).

TARGETING DNA'S MAJOR GROOVE

We reasoned that synthetic agents that targeted the major groove of DNA with recognition through non-covalent surface motifs (as biomolecular recognition units do) could prove to be powerful new tools. The challenge in designing such agents, however, would lie in large part in generating the size of the molecular surfaces required within a synthetic system. For this we turned to the methodology of supramolecular chemistry, one of the hottest areas of synthetic chemical research, which provides excellent methodologies for designing large synthetic arrays of greater scale and complexity than traditional small molecules.[3]

Supramolecular chemistry draws on nature for inspiration and makes use of self-assembly processes to assemble small building blocks into much larger arrays. Of particular relevance to our goals, much supramolecular research employs metal-ligand interactions to control the supramolecular assembly process. The metal centres impart cationic charge to the resulting array and this is particularly attractive for the design of agents that will recognise anionic DNA, since electrostatic attractions will afford a substantial energetic contribution to (non-covalent) binding.

Metal centres have the additional advantage that they are well known to enhance the polarisation of functional groups, which are co-ordinated to the metal centre. This is used to excellent effect in biological systems to catalyse reactions as seen in the zinc enzymes carbonic anhydrase, carboxypeptidase and liver alcohol dehydrogenase. The same effects alter the polarisation of hydrogen-bond donor or acceptors when nucleic acids are metallated, thereby affecting the strength of the base pairing.

Using metallo-supramolecular assembly thus gives the potential to design cationic DNA binders of large dimensions with polarised H-bonding groups on the outside. Such cationic binders have the potential to switch genes on or off through their binding and thus provide a route to control disease.

SUPRAMOLECULAR CYLINDERS

For our proof-of-concept studies we used simple metal-ligand interactions to assemble supramolecular cylinders ~2nm in length and ~1nm in diameter and examined their DNA binding.[4-7] These cylinders are of very similar dimensions to cylindrical zinc-finger motifs, which are the recognition agents for a number of major groove DNA-binding proteins. The synthetic methodology that we developed to assemble these synthetic cylinders allows multi-gramme (or potentially kilogramme) quantities of cylinders to be prepared rapidly (and inexpensively) in one-pot procedures from commercial starting materials.[8] This ease of preparation not only provides plenty of material for examining the DNA binding properties, but allows us readily to modify the nature of the residues on the external molecular surface of the cylinder and the dimensions of the cylinder. The parent cylinder itself is tetracationic and contains three molecular strands wrapped around two iron(II) metal centres in a triple-helical fashion. It is thus chiral and exists in two enantiomeric forms. As we have subsequently shown, the enantiomers have similar effects but different biological modes of action.

Initial DNA-binding studies were undertaken on the racemic mixture of the cylinders.[6,7] A comparison of the dimensions of our synthetic cylinders immediately reveals that they are too large to fit into the minor groove of B-DNA but fit quite snuggly into the major groove stretching across five base pairs. Indeed, the cylinders are slightly longer than zinc finger motifs which target three base pairs (consequently they are deployed in concert by DNA binding proteins). The tetracationic cylinders bind strongly to B-DNA and cannot be displaced by ethedium bromide (we estimate the binding constant to calf thymus DNA to be in excess of 10^7 M^{-1} in 20 mM aqueous NaCl).[6,7] Spectroscopic experiments (UV-Vis absorption and circular dichroism) confirm that the binding event does not significantly perturb the structure of the cylinder and that the DNA remains in a B conformation.[1]

However, the cylinders do have dramatic effects on the structure of the DNA. Flow linear dichroism experiments[9] revealed that the cylinders were bending the DNA. Atomic force microscopy (AFM) images obtained by our colleague Virtudes Moreno in Barcelona have

revealed this to be an intra-molecular coiling of individual strands of double-helical B-DNA.[6] Such a coiling effect is unprecedented with a synthetic agent and reveals the power of designing agents whose dimensions are closer to those employed by biomolecules. Synthetic DNA condensation agents based on charge neutralisation are well established but lead to aggregation of multiple strands.[7] The behaviour seen with the cylinders is quite different: the coiling takes place with individual pieces of double-stranded DNA, and the AFM images suggest that it is a co-operative process that starts at one end of the DNA. Intramolecular coiling of DNA is the key to DNA packaging into the nucleosome.

SEPARATING ENANTIOMERS

Separation of the two enantiomers of our cylinders proved to be quite a challenge using techniques traditionally applied to resolution of metal complexes. However, the solution proved deceptively simple: we were able to demonstrate that the cylinders could be separated by paper chromatography in an experiment akin to the classroom chromatography experiments where the components of inks can be separated by spotting on paper and eluting with water. The process is readily scaled up by using cellulose as a stationary phase in traditional chromatography columns.[10]

Studies on the resolved enantiomers (termed M ($\Lambda\Lambda$) and P ($\Delta\Delta$)) revealed that although both enantiomers are effective coiling agents, the M enantiomer is substantially more aggresive.[4] NMR experiments with a ten-base pair double-stranded oligonucleotide have been conducted by our colleague Einar Sletten in Bergen, Norway and confirm that this enantiomer is indeed in the major groove as predicted.[4,6] Molecular dynamics simulations, restrained by NMR NOE contacts, reveal the DNA to wrap around the cylinder in a fashion that hints intriguingly at the molecular level nature of this unprecedented coiling process.[5]

POSSIBLE APPLICATIONS

The cylinders are superficially reminiscent of histone proteins which wrap up DNA into tight packages while retaining its B-DNA conformation. Possible applications of such an intra-molecular coiling process might include preventing DNA processing, which would lock away the genetic information encoded in that coiled region. This would enable rogue genes to be suppressed. We are currently assessing the extent to which these molecules may be modified and extended to achieve sequence-specific recognition. Sequence-specific recognition coupled with such a coiling process could be a powerful combination indeed and may lead to entirely new classes of drugs with benefits far beyond those currently available. This could, for example, be turning off cancer genes or genes that code for proteins involved in cancerous growth, attacking viral genetic material to defeat viral infection or, perhaps most ambitiously, suppressing the excess genetic material associated with trisomic disorders (such as Down's syndrome).

The remarkable structural effects that we have seen with the cylinders arise in large part from stepping up in size the synthetic molecule towards that employed in nature. Such effects combined with the major groove targeting that the molecular size permits make supramolecular assembly a fascinating playground for the design of new synthetic agents with which we can start to address the challenges thrown down by the genome project to achieve real medical breakthroughs. ●

The authors thank the Leverhulme Trust, the University of Warwick, COST D20 and the EPSRC lifesciences initiative for their support of this research

REFERENCES AND FOOTNOTES

1. B-DNA is the classical Watson-Crick double-helical DNA structure and is the most common form of DNA in nature.

2. See for example White S, Szewczyk JW, Turner JM, Baird EE, and Dervan PB *Nature* 1998;**391**:468-471.

3. See for example (a) the 1987 Chemistry Nobel Prize lecture from J-M Lehn; Supramolecular chemistry – scope and perspectives; molecules, supermolecules and molecular devices, Lehn J-M *Angew Chem Intl Ed* 1998;**27**:89-112. (b) The special feature issue on supramolecular chemistry and self assembly *Proc Natl Acad Sci, USA* 2002;**99**(8) (available free online at www.pnas.org).

4. Meistermann I, Moreno V, Prieto MJ, Moldrheim E, Sletten E, Khalid S, Rodger PM, Peberdy JC, Isaac CJ, Rodger A and Hannon MJ Intramolecular DNA coiling mediated by metallo-supramolecular cylinders: differential binding of P and M helical enantiomers *Proc Natl Acad Sci, USA* 2002;**99**:5069-5074.

5. Molderheim E, Meistermann I, Rodger A, Hannon MJ and Sletten E Interaction between a DNA oligonucleotide and dinuclear iron(II) supramolecular cylinder; a NMR and molecular dynamics study *J Biol Inorg Chem* 2002;**7**:770-780.

6. Hannon MJ, Moreno V, Prieto MJ, Molderheim E, Sletten E, Meistermann I, Isaac CJ, Sanders KJ and Rodger A Intramolecular DNA coiling mediated by a metallo supramolecular cylinder *Chem Intl Ed Engl*, Angew 2001;**40**:879-884.

7. Rodger A, Sanders KJ, Hannon MJ, I. Meistermann I, Parkinson A, Vilder DS and Haworth IS DNA structure control by polycationic species: polyamines, cobalt ammines and di-metallo transition metal chelates *Chirality* 2000;**12**:221-36.

8. Hannon MJ, Painting CL, Hamblin J, Jackson A and Errington W An inexpensive approach to supramolecular architecture *Chem Commun* 1997;1807-1808.

9. In a flow linear dichroism experiment, DNA is oriented in solution by viscous drag. The absorption of the sample in two perpendicular directions is then probed using plane polarised light. The differential absorptions gives information about the orientation of both the DNA and any molecules bound to it. While conceptually quite straightforward the technique is restricted to just a few laboratories in the world. For further information see Rodger A and Nordén B, Circular and linear dichroism, OUP, Oxford, 1987.

10. Hannon MJ, Meistermann I, Isaac CJ, Blomme C, Aldrich-Wright JR and Rodger A Paper: a cheap yet effective chiral stationary phase for chromatographic resolution of metallo-supramolecular helicates *Chem Commun* 2001;1078-1079.

J|A|C|S COMMUNICATIONS

Publication 88

Published on Web 07/24/2003

Hairpin-Shaped Heterometallic Luminescent Lanthanide Complexes for DNA Intercalative Recognition

Peter B. Glover,[†] Peter R. Ashton,[†] Laura J. Childs,[‡] Alison Rodger,[‡] Michael Kercher,[§] René M. Williams,[§] Luisa De Cola,[§] and Zoe Pikramenou*[,†]

School of Chemistry, The University of Birmingham, Edgbaston B15 2TT, U.K., Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, U.K., and Institute of Molecular Chemistry, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WS Amsterdam, The Netherlands

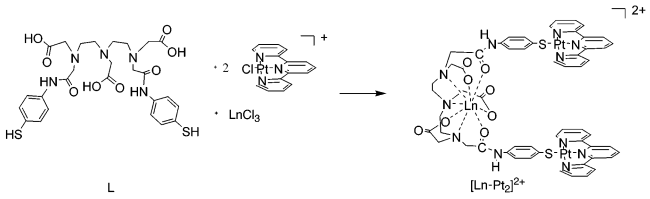
Received December 23, 2002; E-mail: z.pikramenou@bham.ac.uk

Using molecular shape for the functional control of multicomponent systems is one of the challenges in the field of supramolecular chemistry. Interaction of synthetic supramolecular systems with biologically active molecules has been of particular interest to address recognition features important to biomolecular function.[1] We have been interested in the assembly of ligands around lanthanide ions using either steric crowding or replacement of water molecules to control the coordination sphere around the ion.[2] In this paper, we wish to introduce the assembly of lanthanide and transition metal building blocks that lead to metallohairpins. Hairpin-shaped molecules have attracted particular interest for targeting specific DNA sequences.[3] Heterometallic lanthanide complexes have attracted interest in diverse applications ranging from catalytic activity,[4] magnetic devices,[5] and liquid crystalline materials,[6] to sensor design.[7,8] Most of the synthetic strategies have used either transition metal moieties with free sites for lanthanide complexation or ligands with coordination sites designed for metal "induced fit"; there is one example of a covalent bond formation between a lanthanide complex and a metal−porphyrin moiety.[9] The latter complex has been shown to interact with nucleic acids, although most of the lanthanide complexes reported for DNA recognition bear an organic group as an intercalator unit.[10]

In our approach, we use a versatile ligand design that can wrap around a lanthanide, leaving two grafted donor sites for binding of transition metal units. The lanthanide complex is formed using ligand L, a derivative of diethylenetriaminepentaacetic acid[11] with bisamide thiophenol moieties (Scheme 1).

This ligand was developed to provide a hard binding core for lanthanide complexation, via the five oxygens and three nitrogens, and two arms with soft thiol binding sites for binding other metals. Lanthanide complexation orients the ligand to a hairpin-shape with the two thiol sites available for binding. We attached platinum(II) terpyridyl moieties onto these sites using the platinum interaction with aromatic thiolates.[12,13] Platinum(II) terpyridyl moieties are well established as DNA intercalators.[14] The beauty of our approach is the accessibility of the Ln−Pt_2 complexes by self-assembly of the different components in one pot or in a sequential manner, regardless of the order of the addition of the components. Reaction of L, $LnCl_3 \cdot xH_2O$ (Ln = Nd, Eu, La) (the analogous Y(III) complexes were also prepared for NMR spectroscopic characterization), and [Pt(tpy)Cl](CF_3SO_3),[15] where tpy = 2,2′:6′,2″-terpyridine, in a 1:1:2 molar ratio, in methanol under reflux for 2 h yielded exclusively the desired purple Ln−Pt_2 complexes (Scheme 1). Fluorescence titration demonstrates the 1:1 binding of a lanthanide ion and ligand L; see the Supporting Information. The heterometallic complexes were isolated as the PF_6 salts, by precipitation with a methanolic solution of NH_4PF_6 following the addition of water in the reaction mixture and raising of the pH to 5 by NaOH. The same Ln−Pt_2 complexes were obtained independently by two routes: (a) reaction between L and $[Pt(tpy)Cl]^+$ to give L−Pt_2 which was isolated and then reacted with the lanthanide chloride, or (b) by isolation of the complex of L with $LnCl_3$ and subsequent reaction with [Pt(tpy)Cl](CF_3SO_3).

Scheme 1. Preparation of Ln−Pt_2 Metallohairpins



The complexes have been fully characterized and analyzed by spectroscopic methods [selected spectroscopic data **[L−Pt_2]-(PF_6)$_2$** ^{1}H NMR (500 MHz, d_7-DMF, 25 °C) δ 10.3 (br s, NH), 9.06 (m, 6-tpy), 8.79 (m, 3′-tpy), 8.72 (m, 3-tpy), 8.65 (t, 4′-tpy), 8.49 (m, 4-tpy), 7.88 (dd, 5-tpy), 7.56 (m, ArH), 3.4, 3.5, 3.6 (water peak obscures CH_2COO), 3.13 (t, CH_2), 3.00 (t, CH_2); MS (ES^+) *m*/*z* 1460 $\{M - 2(PF_6) - H\}^+$, 731 $\{M\}^{2+}$. Calculated for $C_{56}H_{53}N_{11}O_8S_2Pt_2P_2F_{12}$: C, 38.4; H, 3.1; N, 8.8. Found: C, 38.5; H, 3.2; N 8.6. **[Nd−Pt_2](PF_6)$_2$** MS (MALDI-TOF^+): *m*/*z* 1601 $\{M - 2PF_6\}^+$, (ES^+) *m*/*z* 1748 $\{M - (PF_6)\}^+$, 1601 $\{M - 2(PF_6) - H\}^+$, 801 $\{M - 2(PF_6)\}^{2+}$. Calculated for $C_{56}H_{66}N_{11}O_{16}S_2Pt_2NdP_2F_{12}Na_2Cl_2$: C, 31.2; H, 3.1; N, 7.2. Found: C, 31.3; H, 2.8; N, 7.3. UV−vis (MeOH:H_2O, 4:1): λ_{max} in nm (log ϵ) 515 (3.4), 375 (3.7), 345 (4.4), 330 (4.4), 280 (4.8), 250 (4.8)]. Full assignment of the ^{1}H NMR was achieved by ^{1}H−^{1}H COSY spectroscopy. Electrospray mass spectra for L−Pt_2 and Ln−Pt_2 reveal peaks corresponding to the single charged species $\{M - 2(PF_6) - H\}^+$ and the doubly charged $\{M - 2(PF_6)\}^{2+}$ with the characteristic isotope patterns.

The UV−visible absorption spectra of the Nd−Pt_2 complexes exhibit peaks corresponding to the sum of the bands of the individual components, Pt(tpy)[16] and NdL, with a slight shift of the aromatic thiolate peak and a new broad band centered at 515 nm. This latter band is attributed to a LLCT transition, pπ(RS^-) → π*(tpy) previously observed in $[Pt(tpy)SR]^+$ complexes.[13,17] The appearance of this band is a signature of the formation of the heterometallic complex. The complexes provide two chromophoric components for lanthanide sensitization based on the thiophenol link and the LLCT transition. The emission properties of the Ln−Pt_2 complexes were examined. The Eu−Pt_2 complex does not show any visible emission as expected due to the absorption of the charge-transfer band in this spectral region. However, the Nd−Pt_2 complex shows NIR emission characteristic of the Nd(III) ion at 1060 and

[†] The University of Birmingham.
[‡] University of Warwick.
[§] University of Amsterdam.

10.1021/ja029886s CCC: $25.00

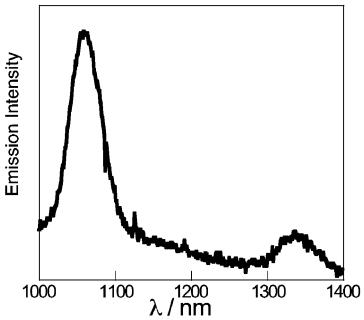


Figure 1. Emission spectrum of $[Nd-Pt_2](PF_6)_2$ in d_6-DMSO, $\lambda_{exc} = 337$ nm.

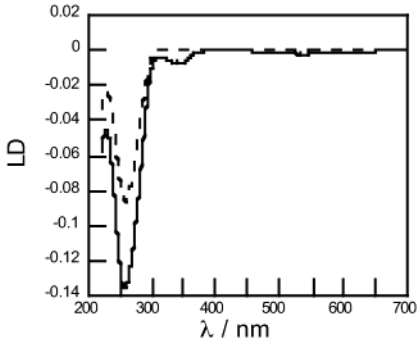


Figure 2. LD spectra of CT-DNA (dashed line), 40:1 CT-DNA:$[Nd-Pt_2]$ complex (solid line); [CT-DNA] = 1000 μM in 10 mM HEPES buffer.

1340 nm corresponding to $^4F_{3/2} \rightarrow {}^4I_{11/2}$, $^4I_{13/2}$ transitions (Figure 1) with a lifetime of 670 ns. The relative quantum yield of the $Nd-Pt_2$ complex upon excitation at 515 nm remains constant in the formation of the complex with calf-thymus DNA. The lanthanide unit acts as a luminescent "reporter" group, the ion being encapsulated by the aminocarboxylate ligating site, which does not interact with DNA. For molecular models, see the Supporting Information.

We have used flow linear dichroism (LD)[18] to detect the binding of the $Nd-Pt_2$ complex to calf-thymus DNA (CT-DNA). The metal complexes will show an LD signal only if bound to DNA. Transitions more parallel than perpendicular to the DNA helix axis will have positive LD signals, while transition moments at an angle of greater than 54° from the helix axis (as is the case for the $\pi-\pi^*$ transitions of planar aromatic intercalators) will result in a negative signal. The LD spectra (Figure 2) show negative signals at 260, 340, and 550 nm.

The 450−610 nm region is composed of a range of transitions that are polarized in the plane of the terpyridines. The LD signal of this region is negative with a very large flat negative LD^r (LD^r = LD/absorbance) signal, indicating that the different terpyridine transitions all lie at the same angle with respect to the DNA helix axis. This indicates the molecule is lying perpendicular to the helix axis. Furthermore, the LD signal of DNA with $Nd-Pt_2$ increases more than the absorbance signal upon $Nd-Pt_2$ binding, showing that the DNA stiffens upon binding. The magnitude of the LD^r signal is significantly greater than that for other bis-intercalators.[19] Control LD experiments with an Ln complex bearing butyl amide arms instead of the −ArSPt(tpy) units show no negative signal for the metal complex, confirming that the neutral lanthanide unit does not bind to DNA. From these results and the shape of the $Nd-Pt_2$ complex, we conclude that the $Nd-Pt_2$ binds to DNA via an intercalative mode that results in significant stiffening of the DNA. This is not only consistent with intercalation but indicative of bis-intercalation of the two −Pt(terpy) units because all of the terpyridyl transition moments are parallel to one another and the terpyridyl region has a LD^r signal more than twice the magnitude of the average DNA base pair even when the DNA is stiffened by the complex binding. The binding constant of $Nd-Pt_2$ with DNA is estimated to be at least an order of magnitude larger than that of ethidium bromide, $K = 9.5 \times 10^6\ M^{-1}$ (see Supporting Information).

We have demonstrated an assembly approach to luminescent heterotrimetallic lanthanide complexes as metallohairpins. The $Ln-Pt_2$ metallohairpins bear intercalating groups that direct the complex to DNA recognition, leading to considerable DNA stiffening with high LD^r values. The lanthanide luminescent unit is "remote" from the negatively charged DNA backbone. Further studies are currently underway to introduce specific recognition sites to target certain DNA sequences.

Acknowledgment. We thank the Royal Society and EPSRC for funding.

Supporting Information Available: Further spectroscopic data of $[Ln-Pt_2]$, UV−vis and LD^r spectra of CT-DNA and CT-DNA + $[Nd-Pt_2]$, emission titration plots of Ln binding to L and competitive binding of $L-Pt_2$ and $Nd-Pt_2$ to DNA, and models of $Nd-Pt_2$ bound to DNA (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Park, H. S.; Lin, Q.; Hamilton, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5105. Hannon, M. J.; Moreno, V.; Prieto, M. J.; Moldrheim, E.; Sletten, E.; Meistermann, I.; Isaac, C. J.; Sanders, K. J.; Rodger, A. *Angew. Chem., Int. Ed.* **2001**, *40*, 880.
(2) Magennis, S. W.; Parsons, S.; Pikramenou, Z. *Chem.-Eur. J.* **2002**, *8*, 5761. Magennis, S. W.; Parsons, S.; Corval, A.; Woollins, J. D.; Pikramenou, Z. *Chem. Soc., Chem. Commun.* **1999**, 61. Magennis, S. W.; Craig, J.; Gardner, A.; Fucassi, F.; Cragg, P. J.; Robertson, N.; Parsons, S.; Pikramenou, Z. *Polyhedron* **2003**, *22*, 745.
(3) Dervan, P. B. *Biorg. Med. Chem.* **2001**, *9*, 2215.
(4) Fischbach, A.; Perdih, F.; Sirsch, P.; Scherer, W.; Anwander, R. *Organometallics* **2002**, *21*, 4569. Hill, M. S.; Hitchcock, P. B. *Angew. Chem.* **2001**, *40*, 4089. Aspinall, H. C.; Greeves, N. *J. Organomet. Chem.* **2002**, *647*, 151. Groeger, H.; Saida, Y.; Sasai, H.; Yamaguchi, K.; Martens, J.; Shibasaki, M. *J. Am. Chem. Soc.* **1998**, *120*, 3089.
(5) Bencini, A.; Benelli, C.; Caneschi, A.; Carlin, R. L.; Dei, A.; Gatteschi, D. *J. Am. Chem. Soc.* **1985**, *107*, 8128. Winpenny, R. E. P. *Chem. Soc. Rev.* **1998**, *27*, 447.
(6) Binnemans, K.; Gorller-Walrand, C. *Chem. Rev.* **2002**, *102*, 2303. Nozary, H.; Piguet, C.; Rivera, J. P.; Tissot, P.; Morgantini, P. Y.; Weber, J.; Bernardinelli, G.; Bunzli, J. C. G.; Deschenaux, R.; Donnio, B.; Guillon, D. *Chem. Mater.* **2002**, *14*, 1075.
(7) Edder, C.; Piguet, C.; Bunzli, J. C. G.; Hopfgartner, G. *Chem.-Eur. J.* **2001**, *7*, 3014.
(8) Klink, S. I.; Keizer, H.; van Veggel, F. *Angew. Chem., Int. Ed.* **2000**, *39*, 4319.
(9) Beeby, A.; Dickins, R. S.; FitzGerald, S.; Govenlock, L. J.; Maupin, C. L.; Parker, D.; Riehl, J. P.; Siligardi, G.; Williams, J. A. G. *Chem. Commun.* **2000**, 1183.
(10) Bobba, G.; Frias, J. C.; Parker, D. *Chem. Commun.* **2002**, 890. Bobba, G.; Kean, S. D.; Parker, D.; Beeby, A.; Baker, G. *J. Chem. Soc., Perkin Trans. 2* **2001**, 1738. Sammes, P. G.; Shek, L.; Watmore, D. *Chem. Commun.* **2000**, 1625. Fu, P. K. L.; Turro, C. *J. Am. Chem. Soc.* **1999**, *121*, 1.
(11) Choppin, G. R.; Schaab, K. M. *Inorg. Chim. Acta* **1996**, *252*, 299.
(12) Jennette, K. W.; Gill, J. T.; Sadownick, J. A.; Lippard, S. J. *J. Am. Chem. Soc.* **1976**, *98*, 6159. Todd, J. A.; Rendina, L. M. *Inorg. Chem.* **2002**, *41*, 3331.
(13) Tzeng, B. C.; Fu, W. F.; Che, C. M.; Chao, H. Y.; Cheung, K. K.; Peng, S. M. *J. Chem. Soc., Dalton Trans.* **1999**, 1017.
(14) Lippard, S. J. *Acc. Chem. Res.* **1978**, *11*, 211. Arena, G.; Monsu Scolaro, L.; Pasternack, R. F.; Romeo, R. *Inorg. Chem.* **1995**, *34*, 2994. Peyratout, C. S.; Aldridge, T. K.; Crites, D. K.; McMillin, D. R. *Inorg. Chem.* **1995**, *34*, 4484. Becker, K.; Herold-Mende, C.; Park, J. J.; Lowe, G.; Schirmer, R. H. *J. Med. Chem.* **2001**, *44*, 2784. McCoubrey, A.; Latham, H. C.; Cook, P. R.; Rodger, A.; Lowe, G. *FEBS Lett.* **1996**, *380*, 73.
(15) Buchner, R.; Cunningham, C. T.; Field, J. S.; Haines, R. J.; McMillin, D. R.; Summerton, G. C. *J. Chem. Soc., Dalton Trans.* **1999**, 711.
(16) Aldridge, T. K.; Stacy, E. M.; McMillin, D. R. *Inorg. Chem.* **1994**, *33*, 722.
(17) Yam, V. W. W.; Tang, R. P. L.; Wong, K. M. C.; Ko, C. C.; Cheung, K. K. *Inorg. Chem.* **2001**, *40*, 571. Howe-Grant, M.; Lippard, S. J. *Inorg. Synth.* **1980**, *20*, 101.
(18) Nordén, B.; Kubista., M.; Kuruscev, T. Q. *Rev. Biophys.* **1992**, *25*, 51. Nordén, B. *Appl. Spectrosc. Rev.* **1978**, *14*, 157.
(19) Kim, H. K.; Kim, J. M.; Kim, S. K.; Rodger, A.; Norden, B. *Biochemistry* **1996**, *35*, 1187.

JA029886S

404 Biophysical Journal Volume 86 January 2004 404–410

Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell

Timothy R. Dafforn,* Jacindra Rajendra,† David J. Halsall,‡ Louise C. Serpell,§¶ and Alison Rodger†

*Department of Biological Sciences, University of Manchester, Manchester M13 9PT, United Kingdom; †Department of Chemistry, University of Warwick, Warwick CV4 7AL, United Kingdom; ‡Department of Clinical Biochemistry and §Department of Haematology, University of Cambridge, Cambridge Institute of Medical Research, Cambridge CB2 2XY, United Kingdom; and ¶Neurobiology Division, Medical Research Council Centre, Laboratory of Molecular Biology, Cambridge CB2 2QH, United Kingdom

ABSTRACT High-resolution structure determination of soluble globular proteins relies heavily on x-ray crystallography techniques. Such an approach is often ineffective for investigations into the structure of fibrous proteins as these proteins generally do not crystallize. Thus investigations into fibrous protein structure have relied on less direct methods such as x-ray fiber diffraction and circular dichroism. Ultraviolet linear dichroism has the potential to provide additional information on the structure of such biomolecular systems. However, existing systems are not optimized for the requirements of fibrous proteins. We have designed and built a low-volume (200 μL), low-wavelength (down to 180 nm), low-pathlength (100 μm), high-alignment flow-alignment system (couette) to perform ultraviolet linear dichroism studies on the fibers formed by a range of biomolecules. The apparatus has been tested using a number of proteins for which longer wavelength linear dichroism spectra had already been measured. The new couette cell has also been used to obtain data on two medically important protein fibers, the all-β-sheet amyloid fibers of the Alzheimer's derived protein Aβ and the long-chain assemblies of α_1-antitrypsin polymers.

INTRODUCTION

Proteins are classified by their structural characteristics to be either globular, membrane bound, or fibrous. Globular proteins are usually highly aqueous soluble whereas membrane proteins are lipid soluble. They provide the majority of the enzymatic, regulatory, and signaling machinery required to maintain a cell. Fibrous proteins in general behave as structural elements both within cells forming the cytoskeleton or in the extracellular environment such as the basement membrane. Fibrous proteins have a range of mechanical properties tailored to their function. Collagen, for example, has, weight for weight, a tensile strength that exceeds that of steel. It is found in connective tissue where such strength is required to maintain tissue integrity under extreme loads (Fratzl et al., 1998; Prockop and Fertala, 1998). The high tensile strength of silk is used by insects for tasks as diverse as catching prey to protection of larval forms. Actin (Otterbein et al., 2001) and myosin, by way of contrast, combine in muscle fibers and perform complex ratchet-like conformational changes that underlie muscle contractions. (Geeves and Holmes, 1999). All these proteins have amino acid sequences that have evolved to form fibers of the correct characteristics. However, in some cases, normally soluble, globular proteins undergo considerable conformational change and polymerize to form β-sheet dominated fibrils. These fibrils are deposited in the tissues leading to diseases collectively known as amyloidoses. Each disease is characterized by protein being deposited as amyloid fibrils. At least 20 different proteins are associated with deposition of such β-sheet fibrils. These include Aβ in Alzheimer's disease, islet amyloid polypeptide in some Type 2 diabetes patients, and prion in the transmissible spongiform encephalopathies. Parkinson's disease and Huntington's disease also involve the deposition of β-sheet fibrils intracellularly. The proteins that form amyloid fibrils do not share any similarities in their native nonfibrillar structure, sequence, or function. Proteins as diverse as lysozyme, immunoglobulin light chain, and apolipoprotein A1 form fibrils in diseased states. The amyloid fibrils formed, however, share morphology and a common core structure (Sunde et al., 1997).

Fibrous proteins constitute the major portion of biomass in the natural world. However, our understanding of the conformations of the polypeptide chains that make up these fibers is limited. This is particularly apparent when compared to our knowledge of globular proteins. It is usually not possible to study the structure of fibrous proteins with conventional techniques used for globular proteins such as single crystal x-ray crystallography or NMR, since fibers are often both insoluble and heterogeneous in length. Instead, electron microscopy (EM) and x-ray fiber diffraction have been utilized to examine the morphology and appearance of the fibers, as well as any repeating structure.

In some cases it has been possible to combine the crystallographic information from a soluble form of a fiber-forming protein with fiber diffraction from a fibrous form (e.g., (Holmes et al., 1990; Rayment et al., 1993) to deduce structural information. The secondary structural elements making up fibers have also been probed using circular dichroism (CD), Fourier transform infrared spectroscopy and solid state NMR. To build up a model of the fibrous protein being studied, several of these techniques are often combined to give information on the spacing of individual elements

Submitted June 2, 2003, and accepted for publication August 21, 2003.

Address reprint requests to Timothy R. Dafforn, Department of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: tim.dafforn@hotmail.com.

0006-3495/04/01/404/07 $2.00

within the fiber, and in some cases they can be combined to gain medium resolution structural information. As the structural study of protein fibers cannot rely on one technique to provide a complete description of the structure, it is all the more important to continue developing new techniques that give unique information on the fibers' construction.

One technique that has the potential to provide information on the structural arrangement of secondary structural elements within the fiber that is not obtainable by the methods so far mentioned is ultraviolet (UV) flow-oriented linear dichroism (LD). The LD signal is produced by measuring the difference in absorbance of linearly polarized light parallel and perpendicular to an orientation direction. LD signals can be positive and negative, positive signals being for transitions whose polarization is along the direction of orientation and negative for those perpendicular to it. These measurements are usually made using a CD spectropolarimeter adapted to produce linear polarized light for LD as opposed to circularly polarized light for CD.

LD has been used extensively to examine the structure of DNA, and in this case, the signal from the π-π^* transitions of the DNA bases is negative as the bases lie more perpendicular than parallel to the DNA helix axis. The use of LD to study protein structure (rather than, e.g., protein DNA complexes) has been limited to early work of Miki and Mihashi (1976), who investigated the signal of nucleotide bound to F-actin, work of Nordh and Nordén (1986) on tubulin, our work on the LD of peptide fibers (Pandya et al., 2000) and our recent work on the LD of proteins bound to liposomes. (Rodger et al., 2002a,b) The early measurements of actin and tubulin LD were restricted to the aromatic regions of the spectrum due to cell and instrument constraints. The LD of the far-UV region where one probes the amide n-π^* ($\sim$220 nm) and π-π^* transitions (208 nm only for the α-helix and $\sim$195 nm for all motifs) has the possibility, like CD, to provide information on the backbone structure of proteins. In particular, LD has the potential to provide detailed information on the orientation of secondary structural elements with respect to the fiber axis. This information has the potential to fill an important gap in the puzzle that surrounds the structures of many fibrous proteins. This was illustrated by our previous work on the n-π^* and the tail of the π-π^* transitions in the peptide fibers, which confirmed that the α-helices ran parallel to the fiber axis. (Pandya et al., 2000) The potential importance of the use of linear dichroism in the study of the conformation of protein fibers has not been realized due to a number of factors including: the limited availability of couette flow cells or other effective orientation equipment; the limited wavelength range of published spectra, which imply that the backbone LD is not really accessible; and the substantial sample requirements of the previously available flow cells. However, our work on α-helical peptide fibers does show the attraction of LD as a technique to study fibers: it gives information that is not readily available by other techniques, making development of an LD cell with characteristics required by biochemistry important.

In practice the most challenging aspect of LD experiments is the fact that the molecules in the sample must be aligned. This has been achieved in the past for proteins using a range of alignment methods, including magnetic fields, stretched films, squeezed gels, and shear flow (Rodger, 1993). We chose to use couette flow as we wished to collect data to as low wavelengths as possible. We present here data on a range of protein fibers that we have obtained with our recently constructed 50 μm annular gap calcium fluoride optics couette flow cell that is illustrated in Fig. 1. To interpret the LD spectra, it is essential to have at least approximate assignments of transition polarizations. Fig. 2 summarized the approximate transition polarizations of the chromophores relevant to the systems studied in this article.

MATERIALS AND METHODS

Sample preparation

Actin was a gift from Dr. Mike Geeves, University of Kent, Canterbury, UK. It was prepared essentially as described by Pardee and Spudich (1982).

Human plasma α_1-antitrypsin was incubated at 60°C for 3 h in 50 mM TrisHCl 100 mM KCl pH 7.6 to produce polymers (Dafforn et al., 1999; Sivasothy et al., 2000; Lomas et al., 1992). Residual monomeric α_1-antitrypsin was then removed by incubation at 95°C for 1 h.

Synthetic A$\beta_{1\text{-}42}$ peptides were obtained from Bachem, St. Helens, UK, and incubated for several months in water at 10 mg/mL at 4°C to produce A$\beta_{1\text{-}42}$ fibrils.

Acid soluble type I collagen was dissolved initially in 100 mM acetic acid. Any insoluble residue was removed by centrifugation at 16,000 $\times$ g for 15 min. The soluble fraction was then dialyzed against 100 mM phosphate buffer at pH 7.4 for 24 h at 4°C.

Spectroscopic experiments

CD measurements were performed using either a Jasco (Tokyo, Japan) J-810 or J-715 spectropolarimeter in 50 mM phosphate buffer pH 7.4 at 20°C using a 0.05 cm pathlength quartz cuvette and averaged over five scans of response time = 1 s.

LD measurements were performed using a Jasco J-715 spectropolarimeter adapted for LD spectroscopy. Samples were placed in 50 mM phosphate buffer pH 7.4 at 20°C and aligned in the light beam using custom-made couette cells (Rodger et al., 2002a,b). The cell (Fig. 1) consists of a cylindrical cross section sleeve with a cylinder mounted centrally with respect to its circular face on a rotating spindle within the sleeve. The spindle center and the center of the circular cross section of the sleeve are aligned so

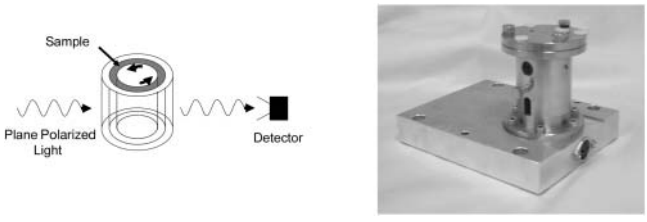


FIGURE 1 Schematic of couette flow LD and picture of the 50 μM annual gap CaF_2 couette flow cell.

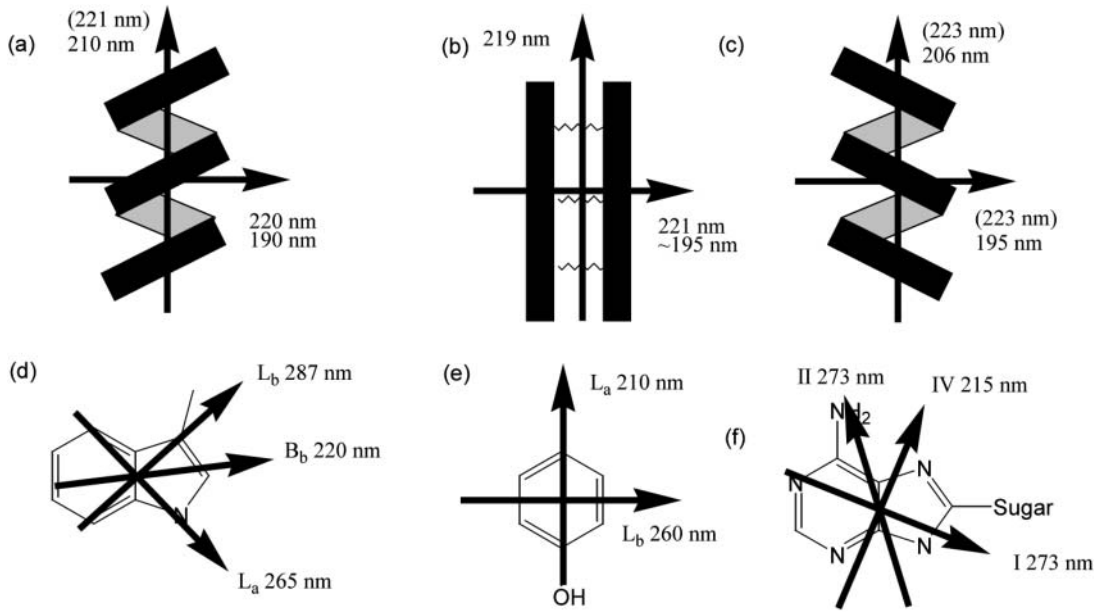


FIGURE 2 The orientations of the various polarization moments in (*a*) α-helix, (*b*) β-sheet, (*c*) poly proline type II, (*d*) tryptophan, (*e*) tyrosine, and (*f*) adenosine chromophores. Arrows indicate transition moment polarizations, and the approximate wavelength maxima of the transitions is indicated near the arrows. In some cases, the common notation for the transition is also given. Wavelengths in parentheses indicate the intensity of this transition is weak.

the cylinder is able to rotate freely within the sleeve. The gap formed between the cylinder and the sleeve is filled with the sample, and upon rotation of the inner cylinder a shear force is induced across the sample. This configuration allows the use of smaller sample volumes (200 μl) than has previously been possible (the previous cell required 2 ml of sample). The light beam is incident radially on the cell and the ones used in this work have two windows on the outer cylinder for the light to pass through. The protein data presented in this paper were all collected with the couette flow cell described in Rodger et al. (2002b). This cell, which is illustrated in Fig. 1, was constructed by Crystal Precision Optics (Rugby, UK). The central rotating cylinder and the entry and exit windows are made of calcium fluoride. This should extend the cell wavelength further into the UV than is obtainable with quartz. Further, the annular gap between the rotating cylinder and windows (see schematic on Fig. 1) is ~50 μm (one-tenth that of existing cells (Rodger, 1993)), which reduces the concentration of nonanalyte molecules (e.g., buffers, Cl^-) in the light beam. Comparison of this cell with previous, long-pathlength cells shows that the increase in the shear force across the 50 μm annular gap compared with the 500 μm annular gap counteracts the expected Beer-Lambert loss of signal due to smaller pathlength and less sample being in the light beam (data not shown). A combination of smaller ''wasted'' sample volume below the windows and the smaller annular gap reduce the required sample volume by approximately an order of magnitude in the CaF_2 cell, which is attractive for studies of biomacromolecules. The voltage applied to the motor that rotates the spindle and hence the CaF_2 cylinder is controlled electronically to allow the sample solution to be maintained with the highest possible degree of alignment without inducing turbulent flow and Taylor vortices.

The shear rate was determined by applying the equations from Nordén et al. (1991) which are based on the work of Taylor (1963). If the inner cylinder rotates:

$$G = \frac{\omega R_i^2 \left[1 + \left(\frac{R_o}{r}\right)^2\right]}{R_i^2 - R_o^2}, \quad (1)$$

where ω is the angular velocity ($d\theta/dt$) and R_i and R_o are the radii of the inner and outer cylinders respectively. This approximates to

$$G \approx \frac{\omega R_o}{R_o - R_i}. \quad (2)$$

If $\omega = 1000$ rpm or alternatively $1000 \times 2\pi/60 = 105\ s^{-1}$; R_o=1.5 cm, 500 μm = 0.05 cm annular gap makes $G \sim 3000\ s^{-1}$.

All spectra were calculated as an average of five measurements.

RESULTS AND DISCUSSION

The 50 μm CaF_2 couette cell was used to measure the LD spectra of four fibrous proteins: collagen type 1; actin, for which LD data exists in the literature thus allowing direct validation of the new cell against older systems; the Alzheimer's related protein A$\beta_{1\text{-}42}$; and α_1-antitrypsin. Before the LD spectrum of each sample was recorded, each was subjected to analysis by CD (Fig. 3). The samples were also examined using negative stain electron microscopy (data not shown). These techniques when combined are able to determine whether the sample is fibrillar and hence amenable to alignment, and also to confirm that the secondary structural content of each fiber is as expected from the literature (see Table 1).

Collagen

Collagen contains peptide bonds in a polyproline type II left-handed extended helix, stabilized by hydrogen bonding interactions with solvent (summarized under ''Other'' in Table 1) with its CD spectrum having a low intensity positive maximum at 220 nm and a much more intense negative maximum at 195 nm (Fig. 3). Collagen fibers are triple PPII helices formed by three individual collagen peptide chains;

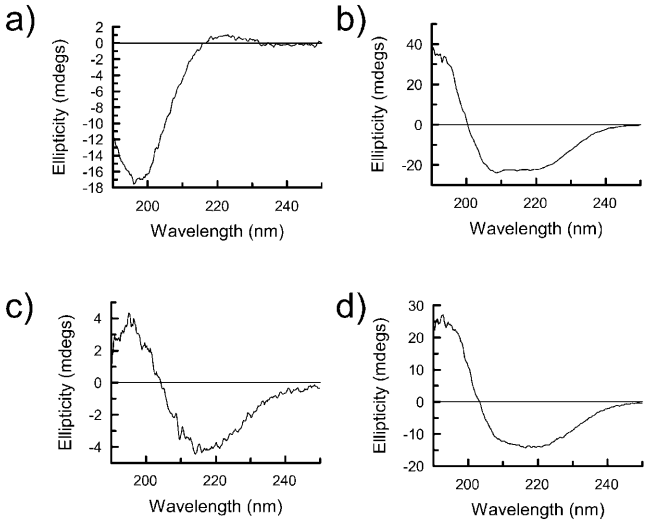


FIGURE 3 CD spectra of (*a*) collagen type I (0.75 mg/ml), (*b*) F-actin (0.75 mg/ml), (*c*) A$\beta_{1\text{-}42}$ (0.32 mg/ml), and (*d*) α_1-antitrypsin (0.72 mg/ml) in a 0.5 mm pathlength cuvette.

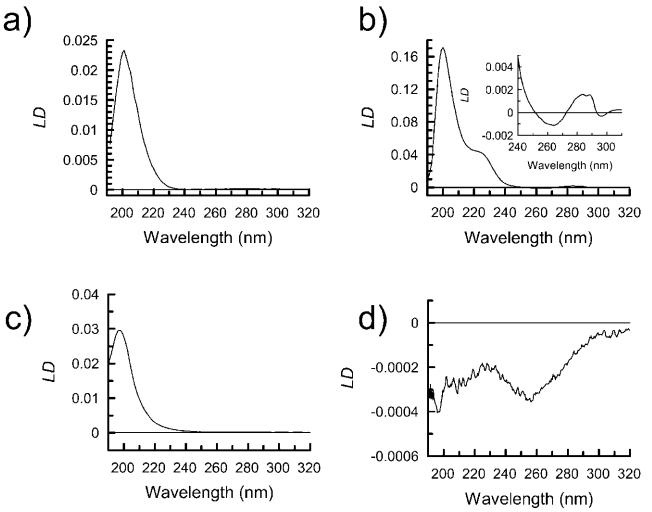


FIGURE 4 LD spectra of the backbone region of (*a*) collagen type I (1.2 mg/ml), (*b*) F-actin (0.05mg/ml) with insert of the aromatic region collected at 0.31mg/ml the concentration, (*c*) A$\beta_{1\text{-}42}$ (0.32mg/ml), and (*d*) α_1-antitrypsin (1.2 mg/ml) in the 50 μm CaF_2 couette cell.

EM shows this sample to contain fibers with average diameters of ~300 Å. The LD spectrum of these fibers (Fig. 4) shows a single positive maximum at 203 nm with a sharp decrease in signal to the low wavelength end of the spectrum, presumably due to a significant negative signal just below 190 nm. There is little LD intensity in the n-π* region at ~220 nm, as is also the case for the polyproline II CD spectrum. The type I collagen used in this study contains a very low percentage of aromatic residues, and hence no LD signal was detectable in the aromatic region. The observed LD requires a transition polarized along the fiber axis at ~206 nm, which is consistent with a transition along the long axis of the poly proline type II helix being aligned with the fiber axis as expected. The transition at lower wavelength is therefore perpendicular to both long axes.

TABLE 1 The secondary structure and aromatic amino acid content of a number of protein fibers

Protein	Number of residues	% Secondary structure content			% Aromatic amino acids		
		α-helix	β-sheet	Other	Trp	Tyr	Phe
Collagen type I	~3000	–	–	100*	–	0.5	1.3
F-actin	337	17†	18†	65†	1.1	4.2	3.2
A$\beta_{1\text{-}42}$	42	–	90*	10*	–	2.3	7.1
α_1-antitrypsin	418	20‡	30‡	50‡	0.7	1.4	6.5

*Estimated from deconvolution of CD spectra.
†Calculated from analysis of x-ray crystallographic structure of monomeric actin (Protein Data Bank code: 1ATN) using DSSP (Kabsch and Sander, 1983).
‡Calculated from analysis of x-ray crystallographic structure of monomeric α_1-antitrypsin (Protein Data Bank code: 1QLP) using DSSP (Kabsch and Sander, 1983).

Actin

The LD signal from proteins, as with CD spectra, can be split into two regions with respect to wavelength and the chromophores present. The near UV region (250 nm–350 nm) contains signals from the aromatic side-chain chromophores, phenylalanine, tyrosine, and tryptophan, as well as certain bound coenzymes such as ATP. The far UV region (250 nm–160 nm) is dominated by backbone chromophores. Actin is a good test sample as it contains a mixture of secondary structural elements as well as a bound cofactor, ATP. The CD spectrum (Fig. 3) of actin was consistent with the presence of both of α- and β-structures. As a result of this, we would expect its LD spectrum to be a sum of the component parts. The measured spectrum for F-actin (Fig. 4) has a significant 220 nm positive LD signal, a negative signal at 210 nm (overlaid by a positive background scattering signal), and a positive maximum at ~200 nm. This requires the average α-helix to be perpendicular to the fiber axis. The 200 nm signal is a combination of the α-helices and β-sheets. With computational theoretical work analogous to the CD calculations of Besley and Hirst (1999) using the matrix method approach (Bayley et al., 1969), it would be possible to determine both the average α-helix and β-sheet orientations.

The significant number of aromatic residues in F-actin (F actin: 1.1% Trp, 3.2% Phe, 4.2% Tyr) as well as the bound ADP (1 molecule per actin subunit), all have the potential to produce an LD signal in the ''aromatic region''. The LD spectrum for F-actin, shown as an inset in Fig. 4 *b*, has significant features in the aromatic region between 250 and 300 nm in accord with the literature (Miki and Mihashi, 1976). The two positive maximum at 285 and 290 nm are due to the ADP (all polymerized actin should have ADP

bound; nonpolymerized is probably ATP), which must be oriented with its plane more parallel than perpendicular to the fiber axis, whereas the 295 nm and 265 nm negative maxima correspond to the tryptophans on average being more perpendicular to the fiber axis.

Actin polymerization monitored by LD

The observation of a significant LD signal for F-actin at 205 nm in the 50 μm CaF_2 LD cell allowed us to follow the polymerization of actin. G-actin polymerization was triggered by the addition of KCl (100 mM) and $MgCl_2$ (2 mM). The kinetics plot obtained (Fig. 5) shows an increase in LD signal over a period of 500 s that conformed to initial lag followed by an increase to saturation. This sort of trace is consistent with those observed by other methods (Millonig et al., 1988; Nishida and Sakai, 1983), and shows the potential of LD for probing fiber formation kinetics—the real advantage of LD being that it only gives a signal for the fibers.

A$\beta_{1\text{-}24}$ and α_1-antitrypsin

Collagen and actin represent protein fibers that are present in high abundance in biological systems. For example, collagen is thought to be the most abundant animal protein in the biosphere. The new cell design opens the application of LD to much less abundant proteins. Two such proteins are the Alzheimer's-related protein A$\beta_{1\text{-}42}$ and the α_1-antitrypsin fiber. Both these proteins have significant importance in the mechanism of disease progression. A$\beta_{1\text{-}42}$ is the major component of the Alzheimer's plaques that form in the brain and is thought to be a causative agent of this form of dementia. The α_1-antitrypsin polymer is the archetypical example of a serpin polymer, and its formation in the liver is a major cause of cirrhosis as well as emphysema. Serpin fibers have also been implicated in a range of other conditions including thrombosis, angioedema (Aulak et al., 1993) and dementia (Davis et al., 1999). Both these proteins are difficult to produce, and the quantities required for LD have in the past precluded their study.

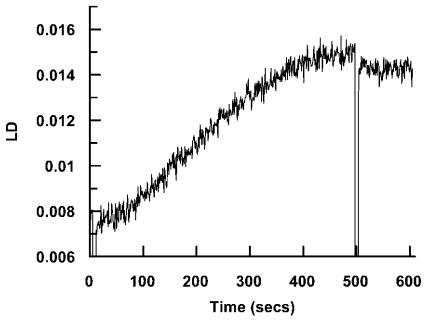


FIGURE 5 Polymerization of actin (0.3 mg/ml) followed by monitoring the LD signal at 205 nm for 500 s. Polymerization was induced by addition of KCl and $MgCl_2$ to a concentration of 100 mM and 2 mM, respectively, at $t = 0$ s. At $t = 500$ s, an excess of KCl and $MgCl_2$ was added (with resulting dilution of the actin and hence immediate decrease of the LD signal) to ensure that the reaction had gone to completion.

A CD spectrum of A$\beta_{1\text{-}42}$ shows a classical β-sheet CD spectrum (Fig. 3) with a negative maximum at 216 nm and a positive maximum at 197 nm. This is in agreement with previous structural studies (Serpell, 2000) that show that these fibers contain predominantly β-structure. Visualization of the samples by EM shows the presence of ordered fibers with an average width of 80 Å. Amyloid fibers typically have a cross β-x-ray fiber diffraction pattern that indicates the β strands run perpendicular to the fiber axis.

The LD spectrum of A$\beta_{1\text{-}42}$ (Fig. 4) is dominated by a single strong positive maximum at ~205 nm with very low signal intensity in the n-π* region of the spectrum (220 nm) that may be due exclusively to background scattering. No evidence for an LD signal within the aromatic region (250−300 nm) was observed despite almost 10% of residues within the A$\beta_{1\text{-}42}$ being either tyrosines or phenylalanines. This indicates that the conformations of the aromatic residues are either not ordered with respect to the β-sheet structure or average to close to the magic angle of 54.7° (Rodger and Nordén, 1997), information that had not previously been available.

X-ray crystal structures of the monomeric form of α_1-antitrypsin show it to have a mixed α/β-structure; this is confirmed to be the case in the polymeric form by observation of an α/β-CD spectrum (Fig. 3). The corresponding LD spectrum (Fig. 4) shows a negative signal at 220 nm and below. The spectrum is qualitatively similar in the backbone region but opposite in sign from that of F-actin, which is also a mixed α/β-protein. The signal-to-noise is much worse, suggesting shorter or less rigid fibers have been formed; this is confirmed by EM studies of the polymers that show them to take the form of ''beads on a string'' rather than the rigid rods of actin. The schematics of Fig. 6 indicating the orientations of the α-helices and β-strands in these two proteins are consistent with these opposite spectra. In both cases, the β-strands are a mix of orientations. The near UV part of the LD spectrum of α_1-antitrypsin shows a single negative minima at 258 nm, which is likely to correspond to the relatively high proportion of phenylalanines (α_1-antitrypsin: 0.7% Trp, 6.5% Phe, 1.4% Tyr) in the protein being well-aligned with respect to one another so that the short axis of the phenol group is perpendicular to the fiber axis.

CONCLUSION

The aim of this study was to show that LD could be used to study expensive biological samples of protein fibers and to show it could be used to aid structural characterization of fibers. The important aspects of lower wavelength range, reduced sample volume, higher optical efficiency, and

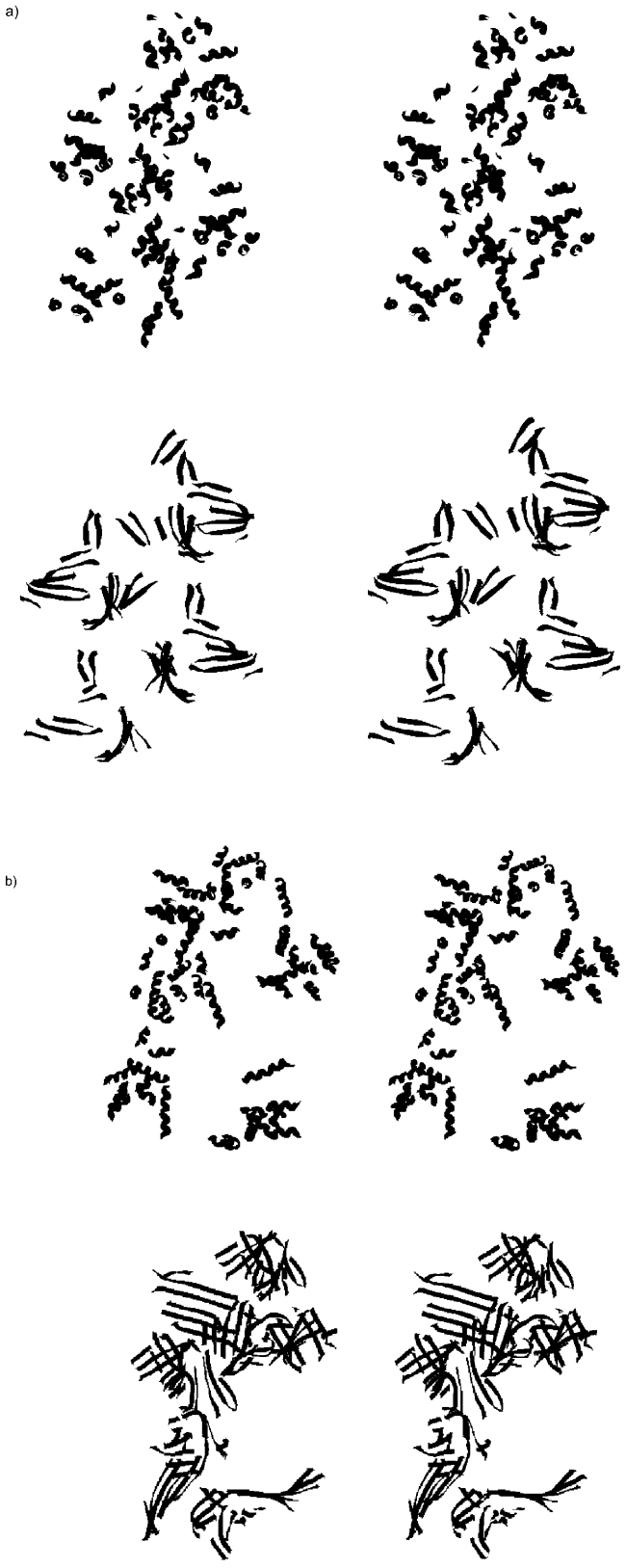


FIGURE 6 Stereo diagrams of positions of secondary structure motifs in (*a*) F-actin and (*b*) α-antitrypsin fibers. In each case, the coils of the top schematic show α-helices and the ribbons of the bottom schematic show β-sheets.

improved alignment have been achieved in the 50 μm CaF_2 couette flow cell. The reduced pathlength of this cell proved to be advantageous to the LD signal due to increased alignment and also reduced absorbance of nonaligned sample components. This has the further advantage that higher protein concentrations can be used in the cell, a factor that can be important in initiation of certain protein oligomerization reactions. Measurement of the LD spectrum of actin showed that our data were consistent with that in the literature and that the cell behaved in a similar manner to those built previously. However, the shortened pathlength enabled collection of data to significantly lower wavelengths. This has implications for the utility of LD for conformational analysis, as significantly more information on secondary structural content is available in the low wavelength regions. The use of the new cell to collect data for two other proteins that are expensive to produce justifies the design of the new cell. In particular, the poor signal-to-noise observed for the α_1-antitrypsin sample would have prevented any information on the structure of the fiber from being obtained in the old long pathlength cell even if enough material could be obtained for such a study.

Some general structure/LD rules can be summarized from our data:

1. Circular dichroism gives a useful handle on the percentages of secondary structure motifs in the fiber.
2. The aromatic region is useful if there is a detectable signal, and where different chromophores have different absorbance wavelengths it should be possible to analyze the data quantitatively to determine their relative orientations. A matrix method-type approach would be most useful (Bayley et al., 1969; Besley and Hirst, 1999).
3. The 220 nm LD is only significant for α-helices, and a negative signal means the average α-helix orientation is more parallel than perpendicular to the fiber axis.
4. For the α-helix, a dip in the spectrum is usually observed at ~210 nm with a sign opposite from the 220 nm n-π^* signal. This is the long wavelength π-π^* component that is unique to the α-helix. Below 200 nm, a signal of the same sign as the n-π^* transition is expected (cf. Fig. 2).
5. The 195−200 nm region is dominated by π-π^* transitions. For poly proline type II motifs, one expects a signal polarized parallel to the helix axis (which in the case of collagen is parallel to the fiber axis, so positive in sign) above 200 nm and one of the opposite sign below 200 nm.
6. β-sheets have one significant band at ~200 nm. If it is positive in sign, then the peptide backbone of the β-strands are more perpendicular than parallel to the fiber axis (cf. Fig. 2).
7. Mixed α/β structures are dominated by the α-helix LD down to ~210 nm. The β-sheet is probably dominant below this, leading to signals at a slightly longer wavelength than the α-helix's lower π-π^* component.

T.D. is a Medical Research Council Career Development Fellow. L.S. is a Wellcome Trust Research Career Development Fellow and thanks the Laboratory of Molecular Biology (Cambridge) for the use of the electron microscope.

A.R. is funded by the Engineering and Physical Sciences Research Council (GR/M91105 and GR/R40869/01).

REFERENCES

Aulak, K. S., E. Eldering, C. E. Hack, Y. P. T. Lubbers, R. A. Harrison, A. Mast, M. Cicardi, and A. E. Davis. 1993. A hinge region mutation in C1-inhibitor (Ala(436)-]Thr) results in nonsubstrate-like behavior and in polymerization of the molecule. *J. Biol. Chem.* 268:18088–18094.

Bayley, P. M., E. B. Nielsen, and A. Schellman. 1969. Rotatory properties of molecules containing two peptide groups: theory. *J. Phys. Chem.* 73:228–243.

Besley, N. A., and J. D. Hirst. 1999. Ab initio study of the of the electronic spectrum of formamide with explicit solvent. *J. Am. Chem. Soc.* 121:9636–9644.

Dafforn, T. R., R. Mahadeva, P. R. Elliott, P. Sivasothy, and D. A. Lomas. 1999. A kinetic mechanism for the polymerization of alpha(1)-antitrypsin. *J. Biol. Chem.* 274:9548–9555.

Davis, R. L., A. E. Shrimpton, P. D. Holohan, C. Bradshaw, D. Feiglin, G. H. Collins, P. Sonderegger, J. Kinter, L. M. Becker, F. Lacbawan, D. Krasnewich, M. Muenke, D. A. Lawrence, M. S. Yerby, C. M. Shaw, D. Gooptu, P. R. Elliott, J. T. Finch, R. W. Carrell, and D. A. Lomas. 1999. Familial dementia caused by polymerization of mutant neuroserpin. *Nature*. 401:376–379.

Fratzl, P., K. Misof, I. Zizak, G. Rapp, H. Amenitsch, and S. Bernstorff. 1998. Fibrillar structure and mechanical properties of collagen. *J. Struct. Biol.* 122:119–122.

Geeves, M. A., and K. C. Holmes. 1999. Structural mechanism of muscle contraction. *Annu. Rev. Biochem.* 68:687–728.

Holmes, K. C., D. Popp, W. Gebhard, and W. Kabsch. 1990. Atomic model of the actin filament. *Nature*. 347:44–49.

Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 22:2577–2637.

Lomas, D. A., D. L. Evans, J. T. Finch, and R. W. Carrell. 1992. The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature*. 357:605–607.

Miki, M., and K. Mihashi. 1976. Fluorescence and flow dichroism of F-actin-epsilon-ADP; the orientation of the admine plane relative to the long axis of F-actin. *Biophys. Chem.* 6:101–106.

Millonig, R., H. Salvo, and U. Aebi. 1988. Probing actin polymerization by intermolecular cross-linking. *J. Cell Biol.* 106:785–796.

Nishida, E., and H. Sakai. 1983. Kinetic analysis of actin polymerization. *J. Biochem.* 93:1011–1020.

Norden, B., C. Elvingson, M. Jonsson, and B. Akerman. 1991. Microscopic behaviour of DNA during electrophoresis: electrophoretic orientation. *Q. Rev. Biophys.* 24:103–164.

Nordh, J. D. J., and B. Nordén. 1986. Flow orientation of brain microtubules studied by linear dichroism. *Eur. Biophys J.* 14:113–122.

Otterbein, L. R., P. Graceffa, and R. Dominguez. 2001. The crystal structure of uncomplexed actin in the ADP state. *Science*. 293:708–711.

Pandya, G. S., M. Sunde, J. R. Thorne, A. Rodger, and D. N. Woolfson. 2000. Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*. 39:8728–8734.

Pardee, J. D., and J. A. Spudich. 1982. Purification of muscle actin. *Methods Cell Biol.* 24:271–289.

Prockop, D. J., and A. Fertala. 1998. The collagen fibril: the almost crystalline structure. *J. Struct. Biol.* 122:111–118.

Rayment, I., H. M. Holden, M. Whittaker, C. B. Yohn, M. Lorenz, K. C. Holmes, and R. A. Milligan. 1993. Structure of the actin-myosin complex and its implications for muscle contraction. *Science*. 261:58–65.

Rodger, A. 1993. Linear dichroism. *Methods Enzymol.* 226:232–258.

Rodger, A., and B. Nordén. 1997. Circular Dichroism and Linear Dichroism. Oxford University Press, Oxford.

Rodger, A., J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J. D. Hirst, A. T. B Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez, and M. J. Hannon. 2002a. Flow oriented linear dichroism to probe protein orientation in membrane environments. *Phys. Chem. Chem. Phys.* 4: 4051–4057.

Rodger, A., J. Rajendra, R. Marrington, R. Mortimer, T. Andrews, J. B. Hirst, A. T. B. Gilbert, R. Marrington, D. Halsall, T. R. Dafforn, M. Ardhammar, B. Nordén, C. A. Woolhead, C. Robinson, T. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez, and M. J. Hannon. 2002b. Flow oriented linear dichroism to probe protein orientation in membrane environments. *In* Biophysical Chemistry: Membranes and Proteins. R. H. Templer, R. Leat, R. J. Leatherbarrow, editors. Royal Society of Chemistry, Cambridge. 3–19.

Serpell, L. C. 2000. Alzheimer's amyloid fibrils: structure and assembly. *Biochim. Biophys. Acta*. 1502:16–30.

Sivasothy, P., T. R. Dafforn, P. G. W. Gettins, and D. A. Lomas. 2000. Pathogenic alpha(1)-antitrypsin polymers are formed by reactive loop-beta-sheet A linkage. *J. Biol. Chem.* 275:33663–33668.

Sunde, M., L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, and C. C. Blake. 1997. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* 273:729–739.

Taylor, E., and W. Cramer. 1963. Birefringence of protein solutions and biological systems. *Biophys. J.* 3:127–141.

Acta Crystallographica Section E
Structure Reports Online
ISSN 1600-5368

Potential injectable contraceptive steroids: testosterone buciclate

Nathaniel W. Alcock,* Karen J. Sanders and Alison Rodger

Department of Chemistry, University of Warwick, Coventry CV4 7AL, England

Correspondence e-mail: n.w.alcock@warwick.ac.uk

Key indicators

Single-crystal X-ray study
T = 180 K
Mean σ(C–C) = 0.007 Å
R factor = 0.073
wR factor = 0.162
Data-to-parameter ratio = 9.1

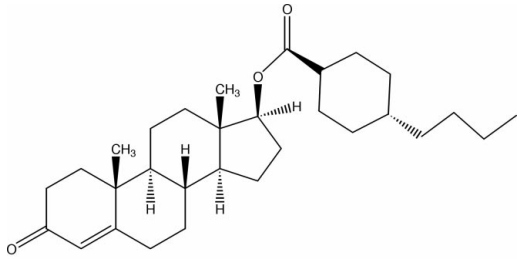
For details of how these key indicators were automatically derived from the article, see http://journals.iucr.org/e.

Received 8 January 2004
Accepted 12 January 2004
Online 14 February 2004

The buciclate ester of testosterone, $C_{30}H_{46}O_3$, whose structure is reported here, was synthesized as part of a World Health Organization programme for the development of injectable contraceptive steroids.

Comment

In the early 1970s, the World Health Organization (WHO) in collaboration with the National Institute of Child Health and Human Development (NICHD) embarked on a steroid ester synthesis programme (Hall *et al.*, 1984). The aim of this programme was to develop long-acting (up to six-month contraceptive activity from a single dose) injectable steroids for both men and women that retained an approximately constant pharmacokinetic profile. An additional aim was to reduce significantly the costs of the products to the public sector. The steroid synthesis programme focused on making new esters, oximes and ethers of steroids with known contraceptive activity. One of the results of this programme was the buciclate ester of testosterone, (I), whose structure is reported here. The dimensions and conformation of the molecule are as expected and match similar compounds already reported (Allen, 2002).



(I)

Experimental

trans-4-*n*-Butylcyclohexanecarboxylic acid (327.4 g, 1.777 *M*, 1.025 equivalents) was dissolved in a mixture of dry dichloromethane (1.5 l, 0.004% *w*/*w* water) and dry dimethylformamide (3 ml, 0.01% *w*/*w* water). Thionyl chloride (226.9 g, 1.907 *M*, 1.1 equivalents) was added to the stirred solution after 5 min at 288–293 K. The reaction mixture was heated gently to 308 K over a period of 0.5 h (a vigorous evolution of gas was observed at 293 K) and then heated at 308–313 K for a further 5 h. The solvent and excess thionyl chloride were removed by rotary evaporation below 313 K under vacuum. Care was taken to protect the resulting acid chloride from moisture at all times. Testosterone (500.0 g, 1.734 *M*) was dissolved in a mixture of dry dichloromethane (2.0 l, 0.004% *w*/*w* water) and dry pyridine (411.4 g, 5.201 *M*, 3.0 equivalents, 0.03% *w*/*w* water). The acid chloride, as prepared above (401.8 g), was dissolved in dry dichloromethane (500 ml, 0.004% *w*/*w* water) and was added dropwise, over a period of 0.5 h, to the testosterone solution at 273–278 K. This mixture was

stirred for a further 2 h at 273–278 K and was then allowed to warm slowly to ambient temperature (293 K) overnight (17 h). The resulting dichloromethane solution was washed with an aqueous 10% citric acid solution (1 l × 3), a saturated aqueous sodium bicarbonate solution (500 ml) and a saturated aqueous sodium chloride solution (500 ml). The resulting dichloromethane solution was dried with magnesium sulfate (300 g) and filtered, and the filtered solids were washed with more dichloromethane (500 ml). The combined dichloromethane filtrates were evaporated to dryness to give the crude testosterone ester (791.1 g). The crude ester was stirred with acetone (1.58 l) for 2.5 h. The resulting solid was filtered and washed with acetone (395 ml). The damp, partially purified, ester was then dissolved in hot methanol (5.5 l) and the hot solution was filtered through glass wool to remove a trace of insoluble material. This solution was stirred as it cooled and crystallized and was held at 283 K for 18 h. The precipitated ester was filtered, washed with methanol (625 ml × 2) and dried *in vacuo* below 313 K. A final purification of the testosterone ester was effected by dissolving it in warm ethanol (2.5 l) and then stirring the mixture at 283 K for 18 h. The recrystallized solid was filtered, washed with ethanol (500 ml) and dried *in vacuo* below 323 K, giving testosterone buciclate (613.5 g, 77.8% yield) as a white crystalline solid.

Crystal data

$C_{30}H_{46}O_3$
M_r = 454.67
Orthorhombic, $P2_12_12_1$
a = 6.2987 (13) Å
b = 10.500 (2) Å
c = 39.680 (8) Å
V = 2624.2 (9) Å^3
Z = 4
D_x = 1.151 Mg m^{-3}
Mo $K\alpha$ radiation
Cell parameters from 4739 reflections
θ = 3–15°
μ = 0.07 mm^{-1}
T = 180 (2) K
Plate, colourless
0.48 × 0.20 × 0.04 mm

Data collection

Siemens SMART diffractometer
ω scans
Absorption correction: multi-scan (*SADABS*; Sheldrick, 1996)
T_{min} = 0.77, T_{max} = 0.96
12 882 measured reflections
2707 independent reflections
1700 reflections with $I > 2\sigma(I)$
R_{int} = 0.124
θ_{max} = 25.0°
$h = -7 \rightarrow 7$
$k = -12 \rightarrow 9$
$l = -44 \rightarrow 47$

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.073
$wR(F^2)$ = 0.162
S = 1.05
2707 reflections
298 parameters
H-atom parameters constrained
$w = 1/[\sigma^2(F_o^2) + (0.0715P)^2]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$ = 0.012
$\Delta\rho_{max}$ = 0.24 e Å^{-3}
$\Delta\rho_{min}$ = −0.21 e Å^{-3}

The temperature of the crystal was controlled using the Oxford Cryosystem Cryostream Cooler (Cosier & Glazer, 1986). H atoms were added at calculated positions and refined using a riding model (C—H distances in the range 0.95–1.00 Å), with U_{iso}(H) equal to 1.2 (or 1.5 for methyl H atoms) times U_{eq} of the carrier atom. Friedel pairs were merged and the absolute configuration was assigned from the known configuration of the steroid moiety of the molecule.

Data collection: *SMART* (Siemens, 1994); cell refinement: *SAINT* (Siemens, 1995); data reduction: *SAINT*; program(s) used to solve structure: *SHELXS*97 (Sheldrick, 1990); program(s) used to refine structure: *SHELXTL* (Sheldrick, 1997); molecular graphics: *SHELXTL*; software used to prepare material for publication: *SHELXTL*.

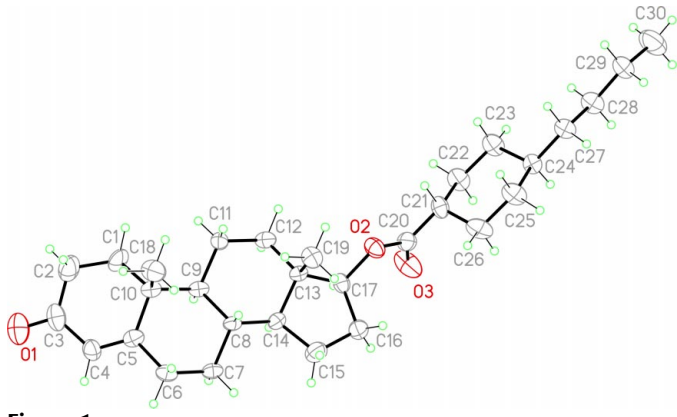


Figure 1
View of the title molecule, showing the atomic numbering scheme. Displacement ellipsoids are drawn at the 50% probability level.

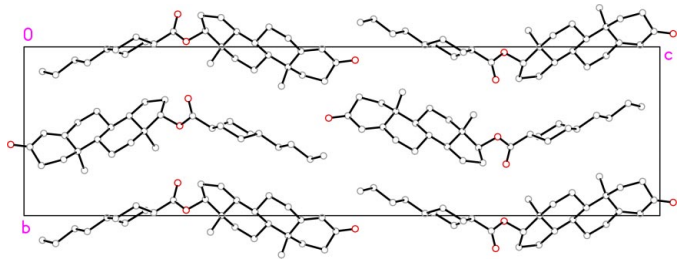


Figure 2
Packing diagram, viewed down the *a* axis. H atoms have been omitted.

We acknowledge the use of the EPSRC's Chemical Database Service at Daresbury Laboratory (Fletcher *et al.*, 1996) for access to the Cambridge Structural Database (Allen, 2002). The EPSRC and Siemens generously supported the purchase of the SMART diffractometer. The Warwick–Kansas collaboration has been supported by *NATO*.

References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Cosier, J. & Glazer, A. M. (1986). *J. Appl. Cryst.* **19**, 105–107.
Fletcher, D. A., McMeeking, R. F. & Parkin, D. (1996). *J. Chem. Inf. Comput. Sci.* **36**, 746–749.
Hall, P. E., Bialy, G., Blye, R. P. & Crabbé, P. (1984). *Long-Acting Contraceptive Delivery Systems*, edited by G. I. Zatuchni, A. Goldsmith, J. D. Shelton & J. J. Sciarra, pp. 190–198. Philadelphia: Harper and Row.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1996). *SADABS*. University of Göttingen, Germany.
Sheldrick, G. M. (1997). *SHELXTL*. Version 5.1. Bruker AXS Inc., Madison, Wisconsin, USA.
Siemens (1994). *SMART*. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.
Siemens (1995). *SAINT*. Version 4. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.

Publication 91

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Enzyme and Microbial Technology 36 (2005) 70–74

ENZYME and MICROBIAL TECHNOLOGY

www.elsevier.com/locate/enzmictec

A study of the secondary structure of *Candida antarctica* lipase B using synchrotron radiation circular dichroism measurements

R.W. McCabe[a], A. Rodger[b], A. Taylor[a,b,*]

[a] *Centre for Materials Science, University of Central Lancashire, Preston PR1 2HE, UK*
[b] *Chemistry Department, University of Warwick, Coventry CV4 7AL, UK*

Received 1 December 2003; accepted 19 April 2004

Abstract

Circular dichroism measurements, using synchrotron radiation, showed that the secondary structure of *Candida antarctica* lipase does not differ significantly when changed from an aqueous to organic solvent environment. Thus, we may conclude that a major conformational change is not the reason for the different product produced by the enzyme when used in organic solvent. Significant changes in the lipase's α-helix content were found at the extremes of pH 4.2 and 9.0; this is in keeping with the permanent loss of activity of the enzyme at such a pH.

Keywords: *Candida antarctica* lipase B; Secondary structure; Circular dichroism

1. Introduction

Over the last 10 years or so there has been much work published on biotransformations using lipases in non-aqueous media [1]. The preferred lipase in many applications is the *Candida antarctica* lipase B, because of its acceptance of a wide variety of substrates, its tolerance of non-aqueous environments and its resistance to thermal deactivation [2–4]. However, it was found by Harffey and coworkers [5] that when the enzyme was used to catalyse the formation of adipate polyesters, a lower molecular weight and higher dispersity polymer was obtained when the reaction was carried out in toluene compared to a solvent free system, adipic acid dissolved in diol substrate without any added solvent

It was known from our kinetic work on the synthesis of adipate polyesters that the concentration of the acid substrate has a significant effect on the activity of the enzyme [6]. Paktar et al. [7] have shown that the optimum pH for *C. antarctica* lipase B is pH 7.0, with a rapid fall in activity below pH 6.0 and above pH 8.0, this effect is usually explained by the importance of the ionisation state of the aspartate and histidine residues of the catalytic triad [8]. However, a permanent decrease in enzyme activity after exposure of the enzyme to low pH was observed which affects adversely the recyclability of the enzyme. A study of the secondary structure of the enzyme creatine kinase using circular dichroism (CD) spectroscopy showed that there was minimal change in the secondary structure between with little loss of activity, after exposure to a pH in between pH 6.0 and 9.0 [9]. Whereas below pH 4.8 and above pH 9.6 there was significant loss of α-helix and of the loss in activity 80% was irreversible.

In this paper, we report an investigation of the secondary structure of *C. antarctica* lipase B in different media to see if there are any changes that might help to explain the different behaviour of the enzyme in the various solvents. In addition, the effect of pH on the secondary structure of the protein is reported. The structure of the *C. antarctica* lipase B enzyme has been determined using X-ray diffraction by a number of workers, notably Uppenberg et al. [10], Cygler and Schrag [11] and Pleiss et al. [12]. Therefore, there was a published foundation on which we could base our studies of the secondary structure of the enzyme.

The use of circular dichroism spectroscopy for the determination of protein secondary structure is quite common; [13] it is, however, used mainly with aqueous systems, but

* Corresponding author. Tel.: +44 1772 893533; fax: +44 1772 892996.
E-mail address: ataylor3@uclan.ac.uk (A. Taylor).

0141-0229/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.enzmictec.2004.04.020

not very much with organic solvents. There are a number of problems that are specific to CD spectroscopy in organic solvents; firstly, it is difficult to obtain solutions of protein in organic solvents and secondly most solvents have a significant absorbance in the UV region [14].

2. Materials and methods

The problem of the significant absorbance of the solvents in the UV region was overcome by using short path length cells and the high photon flux of the Daresbury synchrotron CD beamline 3.1 [15]. This also allowed the collection of data over an extended wavelength range compared to commercial CD spectropolarimeters. The solvents of interest were *n*-hexane and toluene both of which had been used in the enzymatic synthesis of polyesters together with 1,4-butanediol, which although it is a substrate in the polyesterification reaction is also a solvent particularly in the early stages of the reaction.

When using CD spectroscopy to study proteins the choice of buffers available to dissolve the protein is limited by the significant absorbance in the UV of many of the organic buffers. The buffers selected were, 100 mM potassium dihydrogen phosphate to give pH 4.2, 100 mM potassium acetate/acetic acid to give pH 5.0, 100 mM sodium hydrogen phosphate/sodium dihydrogen phosphate to give pH 6.0, 5 mM sodium cacodylate/hydrochloric acid to give pH 7.0 and 100 mM disodium hydrogen phosphate to give pH 9.0.

Protein solutions in the solvents hexane and toluene were prepared by partitioning from a 7 mg dm^{-3} solution in potassium acetate at pH 5.0 into hexane using the method of Dordick [16], whereby the ion pair is formed between dioctyl sulfosuccinate (Aerosol OT) and the enzyme, which then becomes much more hydrophobic and thus soluble in the organic phase. It is not possible to use the Dordick method with the 1,4-butanediol because of the miscibility of water and diol. Therefore, instead of partitioning the protein from the aqueous phase to the organic phase, we used the minimum amount of aqueous buffer necessary to solubilise the protein and added this to the diol. The solutions of *C. antarctica* lipase B were prepared at a concentration of 3 mg dm^{-3} as measured by the UV absorbance at 280 nm using $\varepsilon = 40{,}690\ M^{-1}\ cm^{-1}$ [17].

The secondary structure of the enzyme α-chymotrypsin is well known, it is used as a standard in the CDSSTR program used to predict the secondary structure of proteins from CD spectra [18]. Therefore, in order to validate the proposed methods, the CD spectrum of α-chymotrypsin was determined in both potassium acetate buffer and hexane at a concentration of 3 mg dm^{-3} as measured by the UV absorbance at 280 nm using $\varepsilon = 51{,}350\ M^{-1}\ cm^{-1}$ using the JASCO J715 spectrophotometer between 190 and 270 nm in a 1 mM path length cell.

The CD spectra of *C. antarctica* lipase B were determined over a range of pH and in the solvents hexane, toluene and 1,4-butanediol using the Daresbury SRS Station 3.1, with a 0.01 mm path length cell. The instrument was calibrated using D-10-camphorsulfonic acid in a 0.1 mm cell at the beginning and end of the program. Three scans each of 20 min duration were taken with the aqueous samples and the solvents other than toluene. Because of the unique problems of trying to obtain CD spectra in toluene, which has an absorbance of 1.0 below 284 nm, the sample has to be in the beam for a very long time. 60 min × 20 min scans were obtained in order to determine the toluene background and a further 60 min × 20 min scans to obtain the CD spectra of the protein in toluene. In order to confirm the result, the latter experiment was then repeated using 55 scans with a fresh sample of protein, with similar results.

All solvents used were of spectroscopic grade purity. The α-chymotrypsin (25,193 Da) was purchased from Aldrich, the pure *C. antarctica* lipase B (33,016 Da) was a gift from Novozyme A/S. The Aerosol OT (98% dioctyl sodium sulfosuccinate) was obtained from Aldrich.

3. Results and discussion

The CD spectrum of α-chymotrypsin was measured in aqueous buffer and hexane and the secondary structure distribution determined using the CDSSTR program given in Table 1.

A small decrease in the β-sheet and a 5% increase in turns were observed between the two solvents otherwise the structure is similar to that published. This confirms the work of Yennawar who also found that α-chymotrypsin did not change structure when dissolved in hexane [19].

The CD spectra of the pure *C. antarctica* lipase B (ex Novo) in a number of aqueous buffers ranging from a pH of 4.2 to 9.0 were measured, using the synchrotron radiation source at Daresbury. Three spectra are shown in Fig. 1.

The distribution of secondary structure given by the CDSSTR program are shown in Table 2 and compared to the assignment of secondary structure given by the DSSP/STRIDE algorithm from the E.M.B.L. The DSSP algorithm is a "knowledge based" program, which uses the Brookhaven pdb co-ordinates and assigns each amino acid residue to its most likely form of secondary structure [20].

There is some loss of α-helix at the extremes of the pH range, but the remaining structure does not appear to change significantly over the pH range studied. While the main loss of activity at the extremes of pH is undoubtedly due to the

Table 1
Secondary structure distribution of α-chymotrypsin

Structure	KAc buffer, pH 5.0 (%)	Hexane (%)	CDSSTR standard (%)
α-Helix	8	7	8
β-Strand	38	34	35
Turns	11	16	15
Other	43	43	42

Table 2
Structure of *C. antarctica* lipase B in differing pH buffers

Structure	PO_4 buffer, pH 4.2 (%)	KAc buffer, pH 5.0 (%)	PO_4 buffer, pH 6.0 (%)	Cacod buffer, pH 7.0[a] (%)	PO_4 buffer, pH 9.0 (%)	DSSP (%)
α-Helix	32	38	37	38	30	37
β-Strand	21	20	22	20	25	18
Turns	15	14	11	18	15	26
Other	31	28	30	24	30	19

[a] This result was obtained using conventional CD range 190–270 nm.

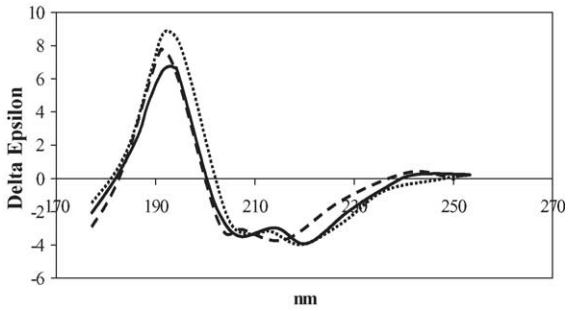


Fig. 1. The CD spectra of *C. antarctica* lipase B at different pH values: pH 5.0 shown as (▪▪▪▪), pH 4.2 shown as (- - - -) and pH 9.0 shown as a black line.

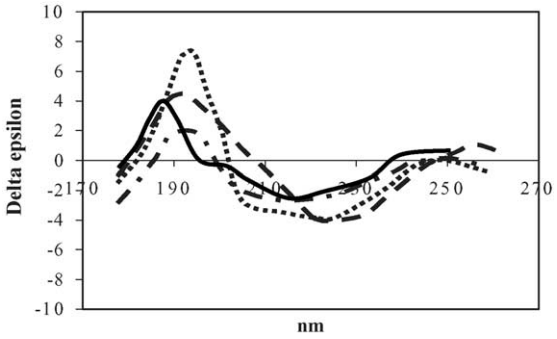


Fig. 2. The CD spectra of *C. antarctica* lipase B in different media. *C. antarctica* lipase B in pH 5.0 aqueous KAc buffer as control (▪▪▪▪). *C. antarctica* lipase B in 1,4-butanediol (- - - -), in hexane (▬▪▬) and in toluene (black line).

effect of the pH on the ionisation of key residues [21], i.e. the Asp 187 and His 224 of the catalytic triad, the permanent loss of activity that occurs after exposure to extremes of pH may be explained by this loss of helix. The spectra obtained after the protein had been transferred from a solution in pH 5.0 potassium acetate buffer to hexane, toluene and 1,4-butanediol are shown in Fig. 2.

The results calculated from the spectra are shown in Table 3 with the structure in the pH 5.0 buffer for comparison.

It must be remembered that the reported structure assignments are based on a fitting program that uses 50 reference spectra from a wide variety of protein structures. In the case of α-chymotrypsin, which is one of the standards, a very close correlation was obtained between the structure obtained from these CD spectra and the published secondary structure. Whereas *C. antarctica* lipase B does not have any proteins of close homology among the reference spectra, therefore, when comparing spectra obtained in solvent compared to reference spectra obtained in aqueous buffers a similar degree of accuracy cannot be expected. The measurement of the distinct CD of the α-helix at 208 nm, however, gives a reasonable confidence in the amount of α-helix quoted, however other structures are not so easy to determine and an error of ±5% may be expected. All spectra were fitted giving a correlation R^2 of 0.95 (rms 0.4) or better, nevertheless the structures other than α-helix should be taken as a guide only.

A number of authors have reported that when the secondary structures of enzymes in solvent are determined by CD spectroscopy there does not appear to be much difference between the structure in solvent and the structure in aqueous buffer [22]. While one might expect the protein to unfold in most organic solvents because of its hydrophobic inner, the stability of the protein may be increased in hydrophobic solvents because of the reduced mobility of the protein chain in the absence of water [23]. This certainly appears to be the case with α-chymotrypsin, as may be seen from the results in Table 1. It also appears to be true for *C. antarctica* lipase B when partitioned from buffer into hexane as shown in Table 3. While this structure was obtained by SRS CD, the result was confirmed subsequently by conventional CD.

The spectrum obtained for the enzyme in toluene (Fig. 3), therefore, is somewhat surprising. Although there was no appreciable drift in the spectrum over the 20 h the specimen was in the instrument (the temperature was kept constant at 25 °C), none the less, the protein might have been affected

Table 3
Structure of *C. antarctica* lipase B in solvents

Structure	Hexane (%)	Toluene observed (%)	Toluene adjusted (%)	1,4-Butanediol (%)	pH 7.0 buffer (%)
α-Helix	35	8	33	33	38
β-Strand	30	31	30	33	20
Turns	15	32	15	14	18
Other	20	29	22	20	24

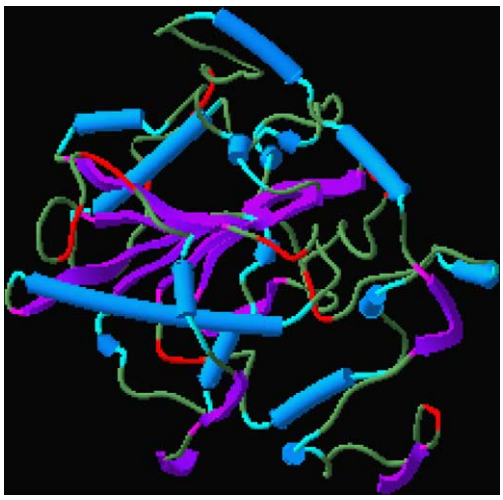


Fig. 3. *C. antarctica* lipase B modelled in Sculpt using the co-ordinates in pdb file 1tca, showing the secondary structure as follows: α-helix ( ), β-sheet ( ), turns ( ) and other ( ).

by the beam. All the experiments were noisy because of the very low levels of light passing through the sample, however, when the spectra from the first 10 scans were averaged and compared to the average of the last 10 scans, no significant difference was seen. Therefore, it is highly unlikely that the sample had deteriorated over time due to the effect of radiation or heat. It can be seen from the spectrum that a bathytopic shift appears to have occurred, with the principal peak moving to a shorter wavelength than the typical helix peak absorbance. The CDSSTR program only assigns helix to the CD absorbance at 208 nm, therefore if for whatever reason the helix absorbance has been shifted then CDSSTR will give an erroneous result. If for some reason the CD spectrum for *C. antarctica* lipase B dissolved in toluene had shifted towards the far UV, as it appears from the spectrum, then the protein has not actually lost α-helix content. This interpretation would be more consistent with the function of the enzyme in toluene. As the shape of the CD spectrum appeared in other respects normal, it was adjusted by adding 10 nm to the wavelength axis and then re-analysed, the results after repositioning of the spectrum are shown in Table 3, column 4. This is almost certainly the first time that a protein structure has been determined in toluene and it could well be the last because it is not easy to justify 3 days of synchrotron time for one experiment.

The extensive α-helix structure of *C. antarctica* lipase B can be seen in the model shown in Fig. 3. The protein has been modelled using the Sculpt program based on the co-ordinates in pdb file 1tca. The active site consisting of the residues Ser 105, Asp 187 and His 224, is labelled and typical of most lipases is on a turn between a short α-helix and a β-sheet, the α–β hydrolase fold. An extensive loss of helix not surprisingly affects the activity of the enzyme, however very long helices in proteins are known to be metastable so that some loss of helix away from the active site may not have a major effect on the activity of the enzyme.

4. Conclusions

It would appear that the secondary structure of *C. antarctica* lipase B in aqueous buffer at pH 5.0–7.0 is very close to the natural structure determined by X-ray crystallography, the isoelectric point, p*I*, of the enzyme being 6.0 [24]. There is a significant loss of α-helix content at extremes of pH, which may well explain the permanent loss of activity, after exposure of the enzyme to pH of 4.2 and 9.0. In hexane and possibly, in the case of the lipase, toluene there appears to be little difference in the secondary structure of either α-chymotrypsin or *C. antarctica* lipase B to that obtained in aqueous buffer, thus confirming the work of Yennawar who also found that α-chymotrypsin did not change structure when dissolved in hexane. Diols such as 1,5-pentanediol are used in protein folding experiments, where they are added to the aqueous medium to enhance the formation of the correct protein structure. Therefore, it is not surprising that in the 1,4-butanediol there appears to be no significant change to the secondary structure of the protein, certainly the enzyme is a highly active esterification catalyst when the acid is dissolved in an excess of the diol. Thus, it appears that any differences in the functionality of the enzyme and the reaction products obtained in toluene and hexane are unlikely to be caused by major changes in the secondary structure of the enzyme.

Acknowledgments

The authors thank Dr. D. Clarke of Daresbury SRS for much advice and help with SRS CD experiments; Dr. Gareth Jones of Daresbury SRS for the extra beamtime necessary to complete this work and Dr. S Paktar of Novozyme A/S for advice and the gift of the pure *C. antarctica* lipase B. The work was supported financially by Baxenden Chemicals Ltd. and the BBSRC with the grant of 1 day's beam time.

References

[1] Roberts SM. Preparative biotransformations: the employment of enzymes and whole cells in synthetic organic chemistry. J Chem Soc, Perkin Trans 1998;1:157–69.
[2] Anderson AE, Larsson KM, Kirk O. One biocatalyst: many applications. The use of *Candida antarctica* lipase B in organic synthesis. Biocatal Biotransform 1998;16:181–204.
[3] Cordova A, Iverson T, Hult K. Lipase catalysed formation of macrocycles by the ring opening polymerisation of ε-caprolactone. Polymer 1998;39(25):6519–24.
[4] Degn P, Pedersen LH, Duus J, Zimmerman W. Lipase catalysed synthesis of glucose fatty acid esters in *t*-butanol. Biotechnol Lett 1999;21:275–80.
[5] Binns F, Harffey P, Roberts SM, Taylor A. Studies leading to the large scale synthesis of polyesters using enzymes. J Chem Soc, Perkin Trans 1999;1:2671–6.
[6] Taylor A. Mechanistic studies of the enzymatic synthesis of polyesters. Ph.D. thesis, University of Central Lancashire; 2001.
[7] Paktar SA, Bjorkling F, Zyndel M, Schulein M, Svendsen A, Gormsen E. Purification of two lipases from *Candida antarctica* and

their inhibition by various inhibitors. Indian J Chem 1993;32B:76–80.
[8] Xie D, Gulnik S, Collins L, Gustchina E, Suvorov L, Erickson JW. Dissection of the pH dependance of inhibitor binding energetics for an aspartic protease: direct measurement of the protonation states of the catalytic aspartic acid residues. Biochemistry 1997;36:16166–72.
[9] Raimbault C, Couthon F, Buchet R. Effects of pH and KCl on the conformations of creatine kinase from rabbit muscle. Eur J Biochem 1995;234:570–8.
[10] Uppenberg J, Hansen MT, Paktar SA, Jones TA. The sequence and crystal structure determination of two forms of lipase B from *Candida antarctica*. Structure 1994;2(4):293–308.
[11] Cygler M, Schrag JD. Structure as a basis for understanding interfacial properties of lipases. Methods Enzymol 1997;284:3–27.
[12] Pleiss J, Fischer M, Schmid RD. Anatomy of lipase binding sites: the scissile fatty acid binding site. Chem Phys Lipids 1998;93:67–80.
[13] Rodger A, Norden B. Circular and linear dichroism. O.U.P.; 1996. p. 15–32.
[14] Lide DR. Handbook of chemistry and physics. 76th ed CRC Press; 1998.
[15] Wallace BA. Synchrotron radiation circular-dichroism spectroscopy as a tool for investigating protein structures. J Synchroton Radiat 2000;7:289–95.
[16] Paradkar VM, Dordick JS. Mechanism of extraction of chymotrypsin into isooctane at very low concentrations of Aerosol OT in the absence of reversed micelles. Biotech Bioeng 1994;43:529–40.
[17] Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 1989;182:319–26.
[18] Hennessey JP, Curtis Johnson W. Information content in the circular dichroism of proteins. Biochemistry 1981;20:1085–94.
[19] Yennawar NH, Yennawar HP, Farber GK. X-ray crystal structure of γ-chymotrypsin in hexane. Biochemistry 1994;33:7326–36.
[20] Frishman D, Argos P. Knowledge based secondary structure assignment. Proteins Struct Funct Genet 1995;23:566–79.
[21] Petersen MTN, Fojan P, Petersen SB. How do lipases and esterases work: the electrostatic contribution. J Biotech 2001;85:115–47.
[22] Dordick JS, Paradkar VM. Aqueous like activity of α-chymotrypsin dissolved in nearly anhydrous organic solvent. J Am Chem Soc 1994;116:5009–10.
[23] Desai UR, Klibanov AM. Assessing the structural integrity of a lyophilised protein in organic solvents. J Am Chem Soc 1995;117:3940–5.
[24] Private communication, S.A. Paktar, Novo A/S.

Micro-Volume Couette Flow Sample Orientation for Absorbance and Fluorescence Linear Dichroism

Rachel Marrington,* Timothy R. Dafforn,† David J. Halsall,‡ and Alison Rodger*
*Department of Chemistry, University of Warwick, Coventry, CV4 7AL, United Kingdom; †Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom; and ‡Department of Clinical Biochemistry, Addenbrooke's Hospital, Cambridge, CB2 2QQ, United Kingdom

ABSTRACT Linear dichroism (LD) can be used to study the alignment of absorbing chromophores within long molecules. In particular, Couette flow LD has been used to good effect in probing ligand binding to DNA and to fibrous proteins. This technique has been previously limited by large sample requirements. Here we report the design and application of a new micro-volume Couette flow cell that significantly enhances the potential applications of flow LD spectroscopy by reducing the sample requirements for flow linear dichroism to 25 μL (with concentrations such that the absorbance maximum of the sample in a 1-cm pathlength cuvette is ~1). The micro-volume Couette cell has also enabled the measurement of fluorescence-detected Couette flow linear dichroism. This new technique enables the orientation of fluorescent ligands to be probed even when their electronic transitions overlap with those of the macromolecule and conversely. The potential of flow-oriented fluorescence dichroism and application of the micro-volume Couette LD cell are illustrated by the collection of data for DNA with minor groove and intercalating ligands: DAPI, Hoechst, and ethidium bromide. As with conventional fluorescence, improved sensitivity compared with absorbance LD is to be expected after instrumentation optimization.

INTRODUCTION

Linear dichroism (LD) is a phenomenon caused by the differential absorption of light polarized parallel and perpendicular to an orientation axis. It has a long history and a wide range of applications, all requiring a means of orienting the sample. For biological molecules, where the sample has to be hydrated, the most versatile sample orientation method has proved to be solution phase Couette flow orientation (Bloemendal and Vangrondelle, 1993; Hofricheter and Eaton, 1976; Nordén et al., 1992; Rodger, 1993). The main applications of such flow LD to date have been to DNA and DNA drug systems (Nordén et al., 1992; Wada, 1964; Wada and Kozawa, 1964). Recent applications of LD for structural characterization of biomacromolecular systems include both membrane-bound and fibrous proteins, which can be difficult to study using other methods (Dafforn et al., 2004; Johansson and Davidsson, 1985; Rodger, 1993; Rodger et al., 2002). However, despite recent sample reduction (Rodger et al., 2002), the sample requirements of flow LD experiments to date (200 μL of a sample whose absorbance is ~1 in a 1-cm cuvette) has precluded its use in systems where limited material is available.

Couette cells are derived from the work of Couette (1890) and were first developed into an instrument for flow dichroism by Wada and Kozawa (1964). Sample orientation is effected by viscous drag generated by a flow gradient imposed on solution flowing in a narrow gap between the walls of a rotating cylindrical cell (Fig. 1). Previous studies have used either a rotating inner and fixed outer cylinder (Wada and Kozawa, 1964), or a rotating outer and fixed inner cylinder (Lee and Davidson, 1968; Oriel and Schellman, 1966). The former design offers more flow stability and has been the most popular. The light path through the cell is made of ultraviolet (UV) transparent materials, usually quartz, that allow optimal transmission of light. If the center rotates then it needs to be made of optically transparent material and the stationary outer cylinder requires at least two transparent windows, and vice versa. The incident light beam can be either perpendicular (Wada and Kozawa, 1964) or parallel (Lee and Davidson, 1968) to the axis of rotation. We adopt the perpendicular orientation in all of our applications.

The micro-volume Couette LD cell developed here is based on a quartz capillary (~5-mm outer diameter and ~3-mm inner diameter) and centrally mounted rod (2–2.5-mm outer diameter) inserted into the capillary (Fig. 2). Flow orientation is generated by rotating the outer capillary. The capillary and rod are demountable for easy cleaning and are in principle disposable. In our previous low-volume Couette (Dafforn et al., 2004; Rodger et al., 2002) the annular gap between an outer stationary cylinder and an inner rotating one was reduced to 50 μm (the standard gap is typically 500 μm). The volume required for the 50-μm cell is ~200 μL, compared to 1800 μL or more (with a 1-mm pathlength) in earlier cells. In this article the new micro-volume Couette LD cell reduces the sample volume by a further order of magnitude (with a pathlength of ~0.5 mm) at the same time opening the way to disposable LD sample holders and also flow-oriented fluorescence dichroism. It is this micro-volume 250-μm annular gap Couette LD cell that has been used for all the work reported in this article.

Submitted September 19, 2003, and accepted for publication June 14, 2004.
Address reprint requests to Alison Rodger, E-mail: a.rodger@warwick.ac.uk.

0006-3495/04/09/2002/11 $2.00
doi: 10.1529/biophysj.103.035022

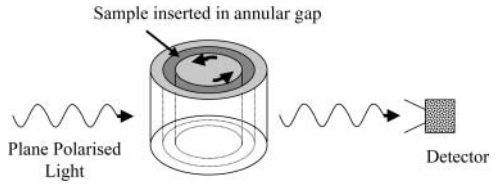


FIGURE 1 Schematic diagram of LD flow Couette cell.

As well as the application of standard LD to new systems (such as liposomes and fibers) a potential area of advance has been the coupling of LD with other techniques. This was first mentioned by Wada (1972), who proposed combination of LD with infrared absorbance fluorescence, and light scattering. Until now, however, fluorescence-detected flow linear dichroism (FDFLD) has not been possible using Couette shear flow due to lack of suitable instrumentation. The micro-volume Couette LD cell offers this possibility because of its small diameter, therefore allowing the collection of photons emitted perpendicular to the incident light direction (Fig. 3). The 90° configuration cell allows the study of fluorophores that exhibit spectral overlap with other species in the solution, so cannot be studied using the standard 180° configuration LD cell. This enhances the capacity of LD for the study of ligand orientation on linear molecules.

LD experiments are usually undertaken using adapted circular dichroism (CD) spectropolarimeters as LD has the same issues as CD with regard to the equivalence of the intensity of the two polarized light beams. In our case we increase the voltage across the photoelastic modulator (PEM) so that it becomes an oscillating half-wave plate with alternating horizontal and vertical polarizations. The key component of the micro-volume Couette LD cell to enable its use in standard CD spectropolarimeters, where light beams are large, is the use of a focusing lens before the capillary housing to ensure the light is only incident on the central part of the unit. The diverging lens effect of the capillary/rod requires another focusing lens after the sample. FDFLD is

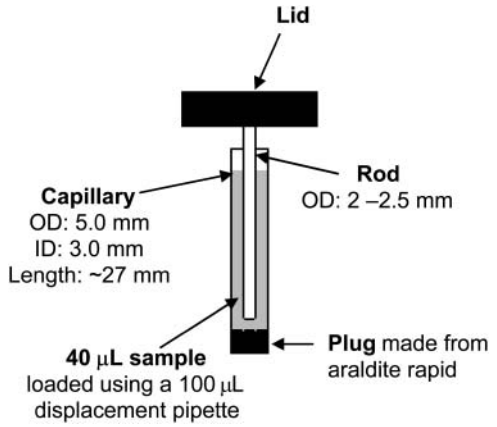


FIGURE 2 Schematic diagram showing the capillary and rod assembly in the micro-volume Couette cell.

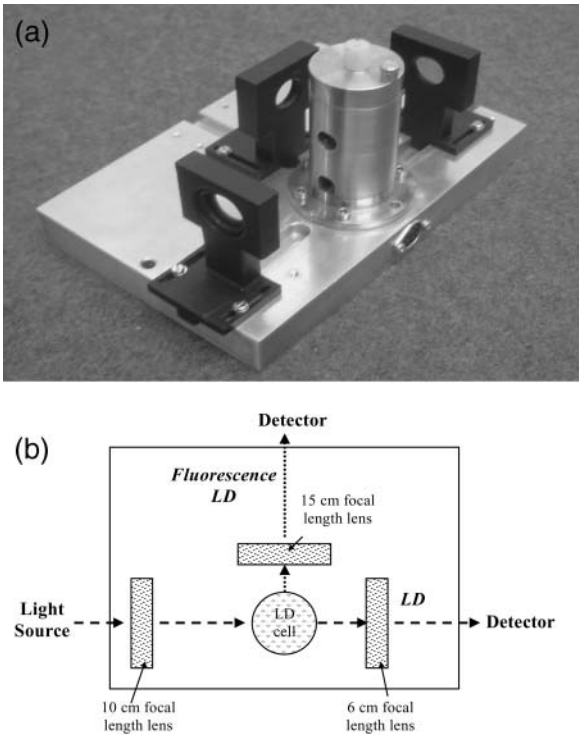


FIGURE 3 (*a*) Photograph of micro-volume Couette flow cell and (*b*) schematic diagram of micro-volume Couette flow cell.

detected by addition of an exit hole at 90° to the incident light and the use of a lens to focus any emitted light onto the photomultiplier tube (PMT), which for the purposes of these experiments was relocated to a position on the back of the CD spectropolarimeter sample compartment at 90° to the incident light (Fig. 3 *b*). This design is not optimized for FDFLD, in particular all scattered photons are detected, resulting in comparatively poor signal/noise ratios. However, as discussed below, we have confirmed that there are no intrinsic problems with the capillary design leading to spectroscopic artifacts.

The LD for a given transition depends on the orientation of the sample (as summarized by the orientation parameter S) and the angle α between the macroscopic orientation axis and the transition moment according to Eq. 1:

$$LD^{\mathrm{r}} = \frac{LD}{A} = \frac{3}{2}S(3\cos^2\alpha - 1), \qquad (1)$$

where LD^{r} is the reduced linear dichroism and A is the absorbance by the sample (in practice, we measure this independently and scale to adjust for pathlength differences between the LD and absorbance experiment).

MATERIALS AND METHODS

Design of micro-volume Couette flow LD cell

A cross section of the micro-volume Couette flow LD cell is shown in Fig. 4. This was designed and built by Crystal Precision Optics, Rugby, UK. The

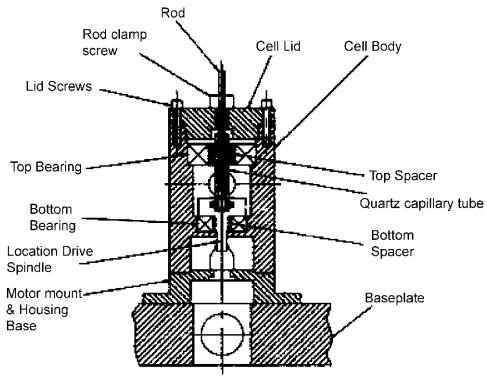


FIGURE 4 General arrangement drawing showing detail of quartz micro-volume Couette flow LD cell.

major innovation in our design is the use of quartz capillaries and rods to hold and orient the sample (Glass Precision Engineering, Leighton Buzzard, UK and Enterprise-Q, Manchester, UK). The capillary units are demountable for removal during cleaning and sample loading. This also provides the possibility for use as a disposable sample holder, an attractive idea in the case of some samples. In this work, the capillaries are sealed at the base with Araldite Rapid (Bostik Findley, Stafford, UK) and held in position in the metal base unit by an ''O''-ring. A quartz rod is held suspended rigidly from the lid and is inserted into the capillary before operation (Fig. 2).

The motor used to drive the capillary is mounted directly below the cell. It is controlled electronically by an EP-603 (0–30 V) power supply, adapted to allow more precise measurements (two decimal places) of applied voltage by the addition of a 10-turn potentiometer. Potential vibration was excluded by using a flexible coupling between the driver spindle and motor. Rotation speeds from 0 to 7000 revolutions per minute (rpm) can be achieved, without inducing turbulent flow. Rotation speeds were ascertained in a calibration experiment by marking the outside of the capillary and measuring the frequency with which the light beam was interrupted using a fast kinetic program on the spectropolarimeter (Fig. 5).

Before collecting the data reported in this article, extensive work was undertaken to demonstrate that different batches of quartz capillaries had no depolarizing effects on the incident light beam. In addition we have shown that the intrinsic LD signal from the capillary is independent of the face of the capillary that is in the light beam. This means that the baseline spectrum for an LD experiment can be collected simply by stopping the rotating capillary and hence the alignment. This is an extremely attractive option and contrasts with our previous practice of emptying and refilling the cell with water. In the case of light-scattering samples, measuring the baseline on the same sample usually produces better baselines. However, the rods are not as uniform as the capillaries and it is essential that they are always inserted into the capillary in the same orientation.

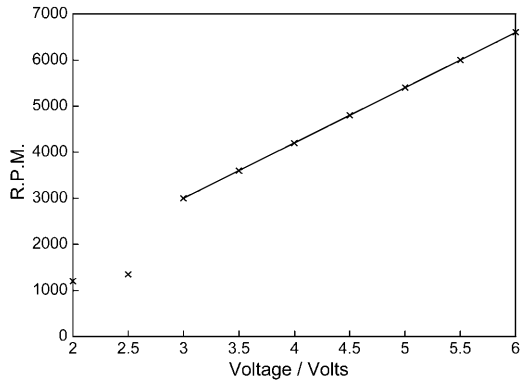


FIGURE 5 A chart to show the relationship between voltage and rpm of the capillary.

Due to the size of the surface area of the capillary and the need to have light incident only on the middle front and back of the solution (otherwise an averaging over samples oriented perpendicular and parallel to the propagation direction would occur) a 2.54-cm diameter × 10-cm focal length lens (Edmund Optics, York, UK) was placed 10 cm in front of the center of the rod. The beam diverges after the capillary unit (in part due to the first lens and in part due to the capillary/rod), so a second lens (2.54-cm diameter × 6-cm focal length) was placed after the sample to focus the light onto the PMT, thereby improving the signal/ noise ratio by maximizing the number of photons reaching the PMT. The focal length of this lens can be varied somewhat without compromising the quality of the data.

The micro-volume 180° Couette flow LD cell was adapted to detect light emitted at 90° by drilling an extra exit hole in the capillary housing at 90° from the path of the incident beam and putting a long focal length lens (2.54-cm diameter × 15-cm focal length) close to this hole to capture and focus maximum light onto the PMT (which is located further from the cell in this configuration).

The minimum volume of sample required depends on the amount of Araldite Rapid used to seal the bottom end of the capillaries because the height of the light beam and the O-ring are fixed. In our standard system a depth of ~4 mm Araldite Rapid was used to seal the capillaries, the rod was lowered to ~1 mm above this when fully assembled, and the total pathlength of the system was ~500 μm. Volumes of at least 25 μL gave the same LD signal, however, higher volumes (e.g., 40 μL) are easier to load without introducing bubbles.

Construction of micro-volume Couette flow LD cell

A major consideration when constructing the micro-volume Couette LD cell was to ensure the rotating quartz capillary and the stationary quartz rod remained parallel. The main body of the micro-volume Couette LD cell is manufactured from laboratory and food industry specification stainless steel (Fig. 4). The bearings and drive spindle were also stainless steel and designed to be dust and water resistant. The drive spindle has a precision-located rubber insert to maintain parallelism and friction drive. The baseplate has been designed specifically for a Jasco CD spectropolarimeter (Tokyo, Japan) with a large sample compartment, but can readily be modified for other CD spectropolarimeters. The base incorporates positioning dowels. These are sufficient to ensure that the cell body is automatically located to line up with the center of the light beam.

The stationary suspended quartz rod is held firmly in position by a cell cap. This has been designed to enable different diameter rods to be installed, allowing the option of changing the annular gap between the inside of the capillary and the outside of the rod without losing the axial parallelism between the capillary and the rod. The base unit of the LD cell has three slots to hold the lenses machined to an accuracy of 10 μm in the linear path and 0.001° in the 90° plane to ensure concentricity in all paths relative to the center line of the rotating capillary. The lens holders were fabricated to ensure that the center of the lens was concentric with the center of the capillary and vertically with the light beam. The optical pathlength of the beam could be adjusted by the linear movement of the lenses along the precision-machined slide ways.

DNA and ligand preparation

Calf thymus DNA, highly lypholized, and the well-characterized DNA groove binders 4′,6-diamidino-2-phenylindole (DAPI) and bisbenzimide H33258 (Hoechst), and ethidium bromide were all supplied by Sigma (Poole, UK). Stock solutions were prepared in water (18.2 MΩ). DNA solutions were prepared by hydration of ct-DNA overnight in deionized

water. Concentrations of ct-DNA were determined spectrophotometrically using the molar absorption coefficient per base of $\varepsilon_{259} = 6600\ M^{-1}\ cm^{-1}$ (Wells et al., 1970). High-purity groove binders/intercalators were used and their concentrations determined by mass using a Sartorius supermicro seven-figure balance (Goettingen, Germany).

Control and validation experiments

The effect of using capillaries as a sample holder for LD and FDFLD was investigated by comparing to reference data collected for LD, CD, and fluorescence. For LD, comparisons against a large-volume Couette flow LD cell (1800 μL) with an inner quartz cylinder that rotates within a steel housing unit with two quartz windows (pathlength 1 mm) was undertaken using ct-DNA solutions at a range of concentrations from 0 to 1000 μM in sodium cacodylate buffer (10 mM, pH 7, Sigma) and NaCl (10 mM, Sigma) (Rodger, 1993). Buffer baseline subtracted LD data were collected from 400 to 190 nm at a data pitch of 0.5 nm, scanning speed of 500 nm min^{-1}, a response of 0.25 s, bandwidth of 2 nm, and averaging over four scans. CD analysis using ammonium camphor sulphonate (0.02% w/v, Jasco) was carried out with the capillary stationary within the baseplate setup and the spectra compared with a 5-mm pathlength quartz cuvette in the normal position for CD analysis. Water baseline corrected CD data were collected between 350 and 185 nm every 0.5 nm, scanning at 200 nm min^{-1}, a response of 0.5 s, bandwidth of 2 nm, and averaging eight scans. Both of these measurements were collected using the Jasco J-715 spectropolarimer in appropriate modes. Fluorescence data were collected using a Perkin Elmer LS50B fluorimeter (Buckinghamshire, UK) not adapted for the use of capillaries, comparing a masked (mimicking the location of the rod) capillary and a 5-mm pathlength quartz cuvette with four polished sides. Aqueous solutions of ct-DNA (317 μM, Sigma), ethidium bromide (30 μM, Sigma) in sodium cacodylate buffer (10 mM, pH 7, Sigma), and NaCl (10 mM, Sigma) were used. Buffer baseline corrected data were collected in excitation mode with the emission wavelength set at 0 nm. Fluorescence data were collected from 600 to 200 nm at a scanning speed of 200 nm min^{-1}, excitation slit widths of 5 nm, and emission slit widths of 2.5 nm. Data were averaged over four scans.

The design of the micro-volume Couette LD cell at present is such that all samples need to be preprepared and loaded into a clean capillary, as it is not possible to remove and replace the rod unit and accurately titrate into the capillary. Validation of the micro-volume Couette LD cell was carried out using ct-DNA (200 μM) comparing repeat analyses in the same capillary and different capillaries. Data were collected using the parameters detailed above under LD control experiments. Analysis of the negative signal at 259 nm yielded information about the errors associated with sample reloading and cell reassembly.

DNA-ethidium bromide binding validation experiments

An application of the micro-volume Couette LD cell is in the analysis of DNA-ligand binding. An example of the binding of ethidium bromide, a classical intercalator, and DNA is shown in this article. The concentration of ct-DNA was held constant at 200 μM and the concentration of ethidium bromide was varied from 5 to 50 μM in 5-μM increments. All solutions were buffered at pH 7.0 using sodium cacodylate buffer (10 mM, Sigma) and contained NaCl (10 mM, Sigma).

LD and fluorescence linear dichroism experiments of DNA-ligand complexes

Applications involving LD and FDFLD of ligands binding to DNA are also reported here. The fluorescent ligands chosen include ethidium bromide, a DNA intercalator with $\alpha \sim 90°$, and two DNA minor groove binders, Hoechst and 4′,6-diamidino-2-phenylindole (DAPI) with $\alpha \sim 45°$ for long-axis polarized transitions and $\alpha \sim 90°$ for short-axis polarized transition. These ligands were expected to provide both positive and negative LD and/or FDFLD signals at a range of different wavelengths. LD and FDFLD experiments of DNA-ligand complexes were performed using calf thymus DNA and groove binder/intercalator solutions of concentration 1000 μM and 50 μM, respectively, thus giving a mixing ratio of 20:1. The solutions were buffered at pH 7 using sodium cacodylate (10 mM, Sigma).

LD and FDFLD data were collected using a Jasco J-715 circular dichroism spectropolarimeter with a large sample compartment that was adapted for LD measurements by turning the PEM into an oscillating half-wave plate. CD artifacts are less likely with this configuration than when a quarter-wave plate is added to an unadapted CD machine. There is no evidence in any spectra of such artifacts. All spectra were taken at room temperature (23°C) using the same quartz capillary and rod with the following parameters: 600–200 nm, data pitch 0.5 nm, scan speed 500 nm min^{-1}, response 0.25 s, bandwidth 2 nm, and averaged over four separate scans. The corrected spectra were calculated by subtracting a baseline from the stationary capillary, when no LD signal would be present.

UV-visible absorbance spectra and fluorescence spectra were recorded using a Jasco V-550 UV-visible spectrophotometer and a Perkin Elmer LS50B fluorimeter, respectively. A 1-mm pathlength quartz cuvette was used to collect UV spectra of ct-DNA and ligands, and a 1-cm pathlength cuvette used for the concentration determination of ct-DNA. Fluorescence excitation spectra were collected in a 5-mm pathlength cuvette with the emission wavelength set to 0 nm, which means all wavelengths emitted and scattered light are collected. This was done to reflect fluorescence data collected in the micro-volume Couette LD system, because the J-715 spectropolarimeter does not have a second monochromator. For fluorescence measurements of the groove binders (DAPI and Hoechst) an excitation slit width of 4 nm and emission slit width of 2.5 nm were used, whereas for ethidium bromide an excitation slit width of 10 nm and emission slit width of 5 nm was used. Data were averaged over four scans at 200 nm min^{-1} scanning speed.

Fluorescence LD control experiments were also carried out using both ct-DNA (1000 μM) and each respective drug (50 μM) at pH 7.0 to ensure that the signal being detected is not due to background scatter or intrinsic properties of the components of the DNA-ligand mixtures.

RESULTS

Verification that capillaries as sample holders do not distort spectra

Control linear dichroism (capillary system versus large-volume Couette flow LD cell with a pathlength of 1 mm), circular dichroism (3-mm inner diameter (i.d.) capillary versus 5-mm pathlength rectangular cuvette) and fluorescence (3-mm i.d. capillary versus 5-mm square fluorescence cuvette) experiments using the capillaries were carried out to show that there are no intrinsic problems with undertaking spectroscopic experiments in the extruded quartz capillaries.

LD spectra of ct-DNA in the range 0–1000 μM measured with the micro-volume Couette LD cell were able to be rescaled so that there was good agreement with data collected in the standard Couette cell across the whole wavelength range down to 200 nm (data not shown). All spectra were consistent with those in the literature (Hofricheter and Eaton, 1976; Houssier et al., 1974; Lee and Davidson, 1968; Nordén et al., 1992; Nordén and Kurucsev, 1994; Simonson and Kubista, 1993; Wada, 1964; Wilson and Schellman, 1978).

CD measurements were undertaken using ammonium camphor sulfonate (0.02% w/v), a standard routinely used for the calibration of CD spectropolarimeters. It yields two signals, a negative signal at 192.5 nm and the other a positive signal at 290.5 nm, with the ratio of the two being 2 (Lowry, 1935; Miles et al., 2003; Takakuwa et al., 1985). The i.d. of the capillaries used is three-fifths that of the reference quartz cuvette. The 3-mm i.d. capillary when rescaled by the appropriate pathlength gave a very good agreement with the reference cuvette data down to 190 nm (data not shown). This shows that the capillary is giving accurate data down to these wavelengths.

Both the LD and CD validation experiments were done with the capillary in the Jasco J-715 CD spectropolarimeters with holders that have been designed for that instrument's baseplate and light beam. With regard to the fluorescence experiments, the controls were undertaken in an LS50B fluorimeter for which we were not able to design a holder that focused the light beam appropriately, therefore the data were of much poorer quality. However, when using ct-DNA (317 μM) and ethidium bromide (30 μM) a masked capillary yielded fluorescence spectra that overlaid with the controls from the same instrument, albeit with less intensity and more noise (data not shown).

Validation of micro-volume Couette flow LD cell

Measurement of a number of individual capillaries established that their pathlength differs by $<1\%$ when capillaries are cut from the same local region of the extruded quartz source material. Multiple LD analyses of the same sample were undertaken, by comparing the signal intensity at 259 nm of a sample of ct-DNA (200 μM) over 15 separate loadings of the capillary. The relative standard deviation (RSD) was 1.1%. An analogous experiment on the same sample without reloading gave an RSD of 0.8%. Thus the errors from realignment and small pathlength variations from sample reloading and cell reassembling contribute an acceptably small amount to overall variation.

Fig. 5 shows the relationship between voltage and rpm using a ct-DNA (200 μM) sample. It can be seen that a minimum voltage of 3 V is required to ensure sufficient rotation of this cell. This value will depend on the motor being used. As the voltage and rpm increases, the intensity of the LD signal also increases. A linear relationship was observed at voltages above 3 V (which fortuitously corresponds to 3000 rpm).

The LD of DNA in the micro-volume Couette LD cell (Fig. 6 *a*) shows a typical DNA LD spectrum: a negative maximum at 259 nm due to π-π* transitions of the DNA bases. Fig. 6 *b* shows a plot of LD_{259} versus DNA concentration at constant rotation speed. These follow the Beer-Lambert law and show the orientation of DNA is not dependent on concentration over this range. The extent of DNA orientation was found to be 10–15% by substituting values for the LD and absorbance and using an average DNA base orientation of 86° (Chou and Johnson, 1993).

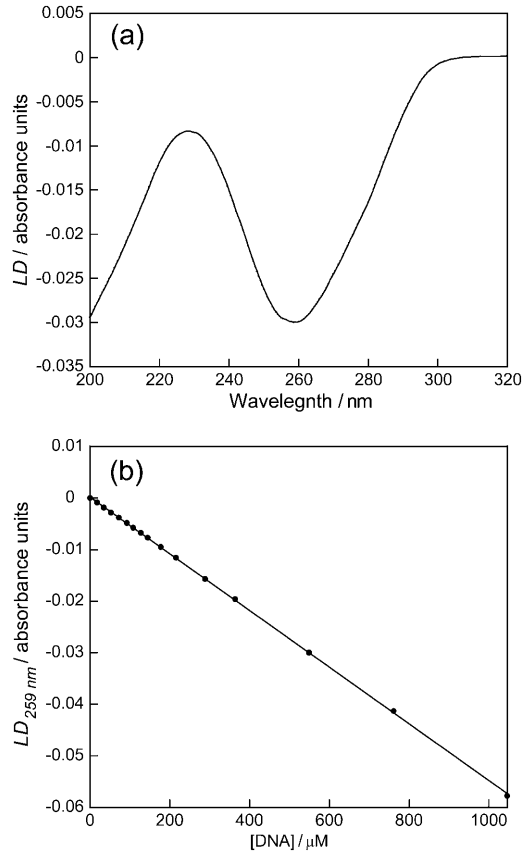


FIGURE 6 (*a*) LD spectrum of ct-DNA (550 μM) in sodium cacodylate buffer (10 mM, pH 7) and NaCl (10 mM); (*b*) LD_{259} versus DNA concentration in a capillary cell at voltage 4 V.

Ethidium bromide-DNA systems

Fig. 7 shows a plot of a titration series of ethidium bromide, a classical DNA intercalator, and DNA. The negative peak at 520 nm is due to π-π* transitions of ethidium bromide; the negative LD signal arises because ethidium bromide intercalates between DNA basepairs in accord with literature (Nordén and Tjerneld, 1976). As the concentration of ethidium bromide is increased, the intensity of this signal increases. Ethidium bromide transitions are also observed below 350 nm in accord with literature expectations.

Fluorescence LD studies on ligands bound to ct-DNA

Before reporting the experimental data from the FDFLD experiments it is important to consider what signal might be expected. In this experiment we are detecting emitted photons, whose number should be proportional to the

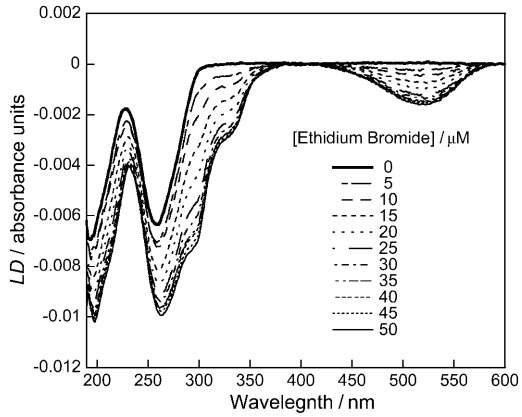


FIGURE 7 LD spectra of DNA (200 μM) and different concentrations of ethidium bromide (0–50 μM) using a sodium cacodylate buffer (10 mM, pH 7) and NaCl (10 mM).

number absorbed (the proportionality constant being the quantum yield). Thus, if we ignore the background scattering contribution to the FDFLD signal, when the sample is being flow oriented, assuming the quantum yield from both polarizations of light are proportional to absorbance with the same proportionality constant, the LD_{90} (the measured FDFLD signal) is expected to be:

$$LD_{90} = \log_{10}\left(\frac{I_{\perp}^{\text{emitted}}}{I_{\parallel}^{\text{emitted}}}\right) = \log_{10}\left(\frac{I_{\perp}^{\text{absorbed}}}{I_{\parallel}^{\text{absorbed}}}\right), \qquad (2)$$

where $I_{\perp}^{\text{emitted}}$ is the intensity of light emitted in the direction of the PMT (which is along the direction of orientation) after absorbance from a perpendicularly polarized light beam and similarly for other terms. By way of contrast, in LD_{180} (the usual configuration) what is detected are the photons that are not absorbed and that proceed straight through the sample:

$$\begin{aligned} LD_{180} &= A_{\parallel} - A_{\perp} \\ &= \log_{10}\left(\frac{I_{\parallel}^{\text{incident}}}{I_{\parallel}^{\text{transmitted}}}\right) - \log_{10}\left(\frac{I_{\perp}^{\text{incident}}}{I_{\perp}^{\text{transmitted}}}\right) \\ &= \log_{10}\left(\frac{I_{\perp}^{\text{transmitted}}}{I_{\parallel}^{\text{transmitted}}}\right), \end{aligned} \qquad (3)$$

assuming the ideal situation where both polarizations of light beams have the same incident intensity. Thus in general qualitative terms the LD_{90} signals are expected to be opposite in sign from those measured in LD_{180}. The more detailed dependence on the angle α the transition moment makes with the orientation direction and the fraction γ of aligned DNA of the LD_{90} measured on a spinning sample (4 V) and completely unoriented (0 V) is (see Appendix):

$$\begin{aligned} LD_{90} &= \left[\lg\frac{(I^{yy} + I^{yx})}{(I^{zy} + I^{zx})}\right]_{4V} - \left[\lg\frac{(I^{yy} + I^{yx})}{(I^{zy} + I^{zx})}\right]_{0V} \\ &\approx -\frac{8\gamma \sin^2\alpha}{3\lg 10}(1 + 2\cos 2\alpha). \end{aligned} \qquad (4)$$

It should be noted that the final line of this equation is based on the assumptions that: i), the sample can be represented as some fully aligned and some unaligned molecules, ii), γ is small, and iii), the DNA does not rotate significantly between excitation and emission. LD_{90} is thus expected to be negative for $0 < \alpha < 60°$ and positive for $60° < \alpha < 90°$. By way of contrast, the change-over point is 54.7° for LD_{180}.

DNA is not expected to have an LD_{90} signal because it is not fluorescent, as is in fact the case. However, upon adding DAPI, Hoechst, and ethidium bromide, the spectra of Figs. 8–10 were observed. Absorbance, LD_{180} and standard excitation fluorescence spectra (with wide open emission slit as described above) are also shown.

DNA-DAPI interactions

The absorption spectra of free and ct-DNA-bound DAPI (Fig. 8 *a*) shows, in accord with literature reports, a red shift of the absorbance maximum by ~15 nm upon binding of DAPI to ct-DNA. There is also a hypochromism of ~20%. The fluorescence excitation spectrum of ct-DNA-bound DAPI with all emitted photons collected is shown in Fig. 8 *b*; this shows a maximum at ~369 nm and a shoulder at ~335 nm. The LD spectra of ct-DNA and ct-DNA in the presence of DAPI are given in Fig. 8 *c*. The ct-DNA shows a characteristic negative peak with maximum at 259 nm. When DAPI is present there is a contribution in this region of the spectrum as seen by a wavelength shift of the peak maximum from 259 nm to 257 nm. There is also a small increase in the ct-DNA LD signal intensity. A broad positive peak centered at ~345 nm is due to the long-axis polarized transition of the groove bound DAPI (where it is expected to bind for AT-rich DNA (Eriksson et al., 1993; Kubista et al., 1987)). There is also a slight negative region at 400 nm due probably to a small population of major groove bound, probably partially intercalated, DAPI at GC-rich regions (Eriksson, 1992; Kim et al., 1993).

The FDFLD spectrum of DAPI bound to ct-DNA is given in Fig. 8 *d*. There is a large negative peak at ~350 nm corresponding to the positive LD_{180} band in Fig. 8 *c*. There is also a smaller negative peak at ~265 nm. This feature corresponds to the 262-nm peak in the DAPI absorbance spectrum that is approximately half the intensity of the one at 347 nm. Thus LD_{90} is able to probe ligand transitions usually masked by the DNA absorbance. It should be noted that the FDFLD intensity will be dependent on the quantum yields of different binding modes and/or different sequences. In addition there is also the possibility of enhanced intensity

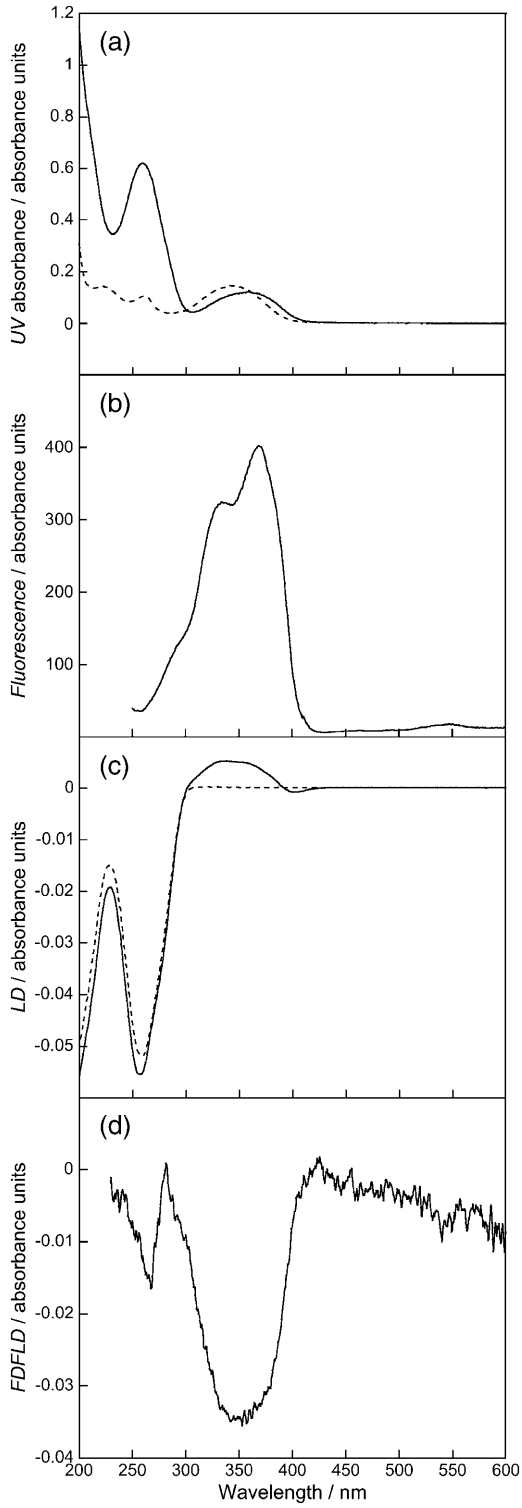


FIGURE 8 Spectra of DNA (1000 μM) and DAPI (50 μM). (*a*) Absorption spectra of DAPI (*dashed line*) and DAPI-ct-DNA (*solid line*), (*b*) fluorescence excitation spectrum of DAPI-ct-DNA with all emitted photons detected, (*c*) LD spectra of ct-DNA (*dashed line*) and DAPI-ct-DNA (*solid line*), and (*d*) FDFLD spectrum of DAPI-ct-DNA. All solutions prepared using a sodium cacodylate buffer (10 mM, pH 7).

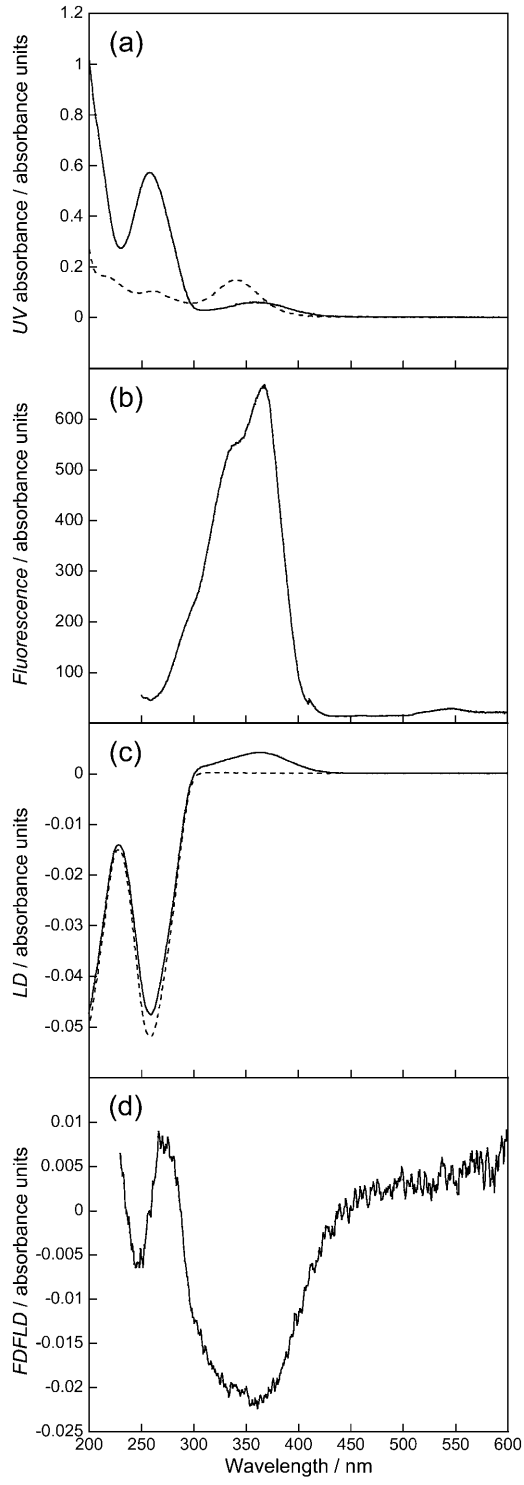


FIGURE 9 Spectra of DNA (1000 μM) and Hoechst (50 μM). (*a*) Absorption spectra of Hoechst (*dashed line*) and Hoechst-ct-DNA (*solid line*), (*b*) fluorescence excitation spectrum of Hoechst-ct-DNA with all emitted photons detected, (*c*) LD spectra of ct-DNA (*dashed line*) and Hoechst-ct-DNA, and (*d*) FDFLD spectrum of Hoechst-ct-DNA. All solutions prepared using a sodium cacodylate buffer (10 mM, pH 7).

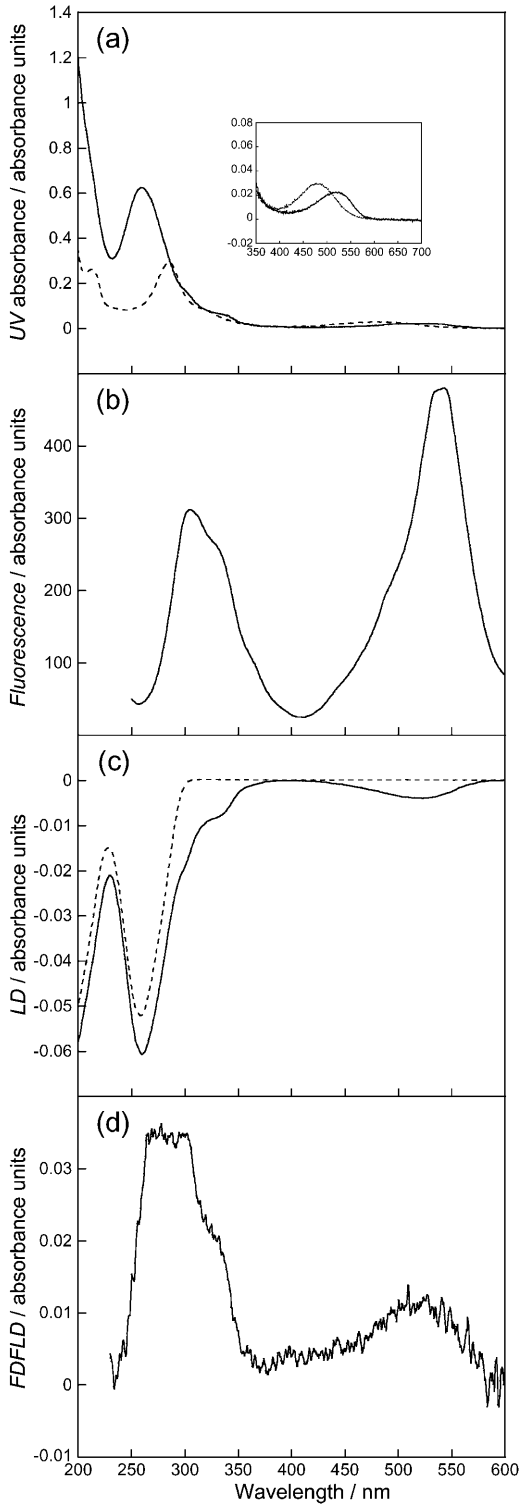


FIGURE 10 DNA (1000 μM) and ethidium bromide (50 μM). (*a*) Absorption spectra of ethidium bromide (*dashed line*) and ethidium bromide-ct-DNA (*solid line*). (*Inset*) Enlarged region of ethidium bromide band. (*b*) Fluorescence excitation spectra with all emitted photons detected, (*c*) LD spectra of ct-DNA (*dashed line*) and ethidium bromide-ct-DNA (*solid line*), and (*d*) FDFLD spectra. All solutions prepared using a sodium cacodylate buffer (10 mM, pH 7).

due to energy transfer from the DNA bases. Given the magnitude of this peak such an effect is not obvious in this spectrum.

DNA-Hoechst interactions

The absorption spectra of free and ct-DNA-bound Hoechst is shown in Fig. 9 *a*. It can be seen that there is red shift of absorbance maximum of ~20 nm upon the binding of Hoechst to ct-DNA. There is also a hypochromism of ~60%. Fig. 9 *b* shows the fluorescence excitation of Hoechst bound to ct-DNA with all emitted photons collected. A peak maximum is seen at ~367 nm with a shoulder at ~338nm. The LD spectra of ct-DNA and ct-DNA in the presence of Hoechst are given in Fig. 9 *c*. The ct-DNA shows a characteristic negative peak at 259 nm. When Hoechst is present there is a decrease in LD signal intensity compared to ct-DNA alone that could be a Hoechst contribution or a reduction in DNA alignment. It has previously been concluded from studies by Moon et al. (1996) that Hoechst binds in the minor groove of AT-rich regions (though the consensus of opinion is that Hoechst binds in the major groove of GC-rich DNA probably in some kind of partially intercalated binding mode; Colson et al., 1995; Moon et al., 1996), and the positive peak ~360 nm is due to the Hoechst bound to ct-DNA along the minor groove.

Fig. 9 *d* shows the FDFLD spectrum of Hoechst bound to ct-DNA. There is a broad negative peak in the 290–440 nm region. This corresponds well with the fluorescence excitation data collected and with the broad positive peak at ~360 nm in the LD spectrum. The positive FDFLD peak at 273 nm is the red-shifted signal of the 260-nm free Hoechst absorbance, consistent with this being a long-axis polarization transition whose LD signal is usually buried under that of the DNA. The superficial similarity of the DAPI and Hoechst FDFLD in the DNA region is misleading as one should note that the baselines slope so the 260-nm region is an FDFLD maximum for DAPI, whereas for Hoechst the same region is an absorbance minimum so not a signal. Similarly the ~270-nm positive peak is in fact zero for DAPI and an FDFLD signal for Hoechst.

DNA-ethidium bromide interactions

Fig. 10 *a* shows the absorption spectra of free and ct-DNA-bound ethidium bromide. The absorption maximum at 525 nm shows a red shift of ~40 nm upon the binding to ct-DNA. There is also a hypochromism of ~25%. The fluorescence excitation spectrum of ethidium bromide bound to ct-DNA with all emitted photons collected is shown in Fig. 10 *b*. A peak maximum at ~300 nm with a shoulder at ~330 nm is observed. A more intense peak is also present at ~530 nm. The LD spectra of ct-DNA and ct-DNA in the presence of ethidium bromide are given in Fig. 10 *c*. The ct-DNA shows a characteristic negative peak at 259 nm.

When ethidium bromide is present there is an increase in LD signal intensity compared to ct-DNA alone. Ethidium bromide is a well-known and well-characterized intercalator. This means that the aromatic rings of ethidium bromide are able to insert between basepairs on the DNA making the DNA more rigid therefore increasing alignment (Nordén and Tjerneld, 1976). There are two negative signals due to the ethidium bromide, one at ~520 nm and one showing as a shoulder and wavelength shift on the DNA signal at ~300 nm.

Fig. 10 *d* shows the FDFLD spectrum of ethidium bromide bound to ct-DNA. There is a broad positive peak with a maximum at 525 nm, corresponding to the negative LD band at that wavelength. There is also a broad positive peak from 250 to 350 nm corresponding to the absorbance peaks in this region. These signals are consistent with an intercalated ethidium bromide and show that it is possible to use absorbance signals that are usually complicated by being masked by DNA spectroscopy. Enhanced signals due to energy transfer from the DNA bases are more likely for an intercalator than for a groove binder. Once again although this possibility cannot be ruled out at this stage, the ethidium FDFLD does not resemble the DNA absorbance shape, so we conclude at least that this is not a dominant effect.

CONCLUSIONS

A micro-volume Couette flow cell has been designed and developed for linear dichroism spectroscopy for applications where sample availability is restricted. This is particularly a relevant issue for biological samples. Flow orientation is achieved by the alignment of linear molecules between a fixed quartz rod and rotating quartz capillary. The capillaries used are optically uniform and have advantages of being cheap and removable, which means that they have the potential of being disposable as well as being significantly easier to clean than other LD cells. Focusing lenses are used to focus the light beam onto the capillary, therefore, maximizing the photon count interacting with the sample. The capillaries do not distort LD, CD, or fluorescence spectra as shown by comparison with data collected using standard sample cells. Method validation using calf thymus DNA and groove binding and intercalating DNA binding ligands has shown sample volumes of ~25 μL are sufficient to obtain reproducible stable LD spectra, although slightly larger volumes reduce the risk of introducing air bubbles into the capillary upon loading. The ct-DNA has been oriented with an orientation parameter of $S = 10$–15% assuming an angle of 86° between the macroscopic orientation axis and the DNA base transition moment.

The capillary design of the micro-volume Couette flow LD cell has also enabled a new technique of fluorescence-detected flow linear dichroism to be undertaken. Exit holes from the cell at both 90° and 180° to the incident light enable measurement of fluorescence LD and normal LD, respectively. In both cases lenses have been placed close to the exit hole to capture the maximum amount of the emitted light and focus this toward the PMT. For measurements of fluorescence LD it was also necessary to move the PMT to be at right angles to the incident light beam.

The FDFLD showed the intrinsic contribution to the LD of the bound ligands whose bands lie under the DNA bands thus showing the potential of FDFLD for probing fluorophores in LD independently from other chromophores. The data are of sufficient quality that it has been possible to probe clearly transitions that are usually hidden by overlapping DNA absorbances. The quality of the fluorescence LD data are less good than the standard LD signal, however, due to the configuration of the spectrometer rather than the cell design. The use of filters or a second monochromator would improve the signal/noise in these experiments. Work is in progress to design a new spectrometer with smaller light beams and both fluorescence and linear dichroism capabilities to significantly improve the quality of the FDFLD data. It will be advantageous for this experiment to record actual differences in intensities of emitted light rather than logarithms of ratios.

APPENDIX

Fluorescence detected flow linear dichroism equations

The shear force of the Couette flow gives a preferential orientation direction to the sample. For simplicity in this preliminary analysis let us assume that a fraction, γ of the DNA is perfectly oriented and a fraction, $1 - \gamma$ is randomly oriented. Due to this experiment design, which uses a Jasco J-715 spectropolarimeter, our data output is the logarithm of the ratio of the signals from the horizontal and vertical polarizations of linearly polarized light. Thus, following a thought process analogous to that derived by Canet et al. (2001) (for fluorescence polarization anisotropy with a photoelastic modulator and a single monochromator with no filters) our LD_{90}, the signal measured when the J-715 instrument is in LD mode and the photomultiplier tube is at 90°, is:

$$LD_{90} = -\lg(I^{zy} + I^{zx}) + \lg(I^{yy} + I^{yx}), \qquad \text{(A1)}$$

where I^{zy} denotes light that is polarized along z with propagation direction along x and that is subsequently detected with polarization along the y direction with propagation direction along z, etc. We may thus write,

$$I^{zy} = k\mu_z^{ex}\mu_z^{ex}\mu_y^{em}\mu_y^{em}, \qquad \text{(A2)}$$

where k depends on light intensity, pathlength, sample concentration, etc. and μ_y^{em} is the y component of the electric dipole transition moment when it emits radiation; similarly *ex* is for excitation.

If we assume that the molecule does not rotate significantly while it is in the excited state (DNA is large so molecules bound to it rotate slowly), then we can drop the *ex* and *em* labels and write

$$\boldsymbol{\mu} = \mu(\sin\alpha\cos\beta,\ \sin\alpha\sin\beta,\ \cos\alpha), \qquad \text{(A3)}$$

where α is the angle between $\boldsymbol{\mu}$ and the z axis and β is the angle between the projection of $\boldsymbol{\mu}$ onto xy plane and x axis then, e.g.,

$$I^{zy} = k\mu^2(\cos^2\alpha\sin^2\alpha\sin^2\beta). \qquad \text{(A4)}$$

A rotational average with respect to β leads to:

$$I^{zy} = k\mu^2\frac{(\cos^2\alpha\sin^2\alpha)}{2} = I^{zx} \tag{A5}$$

$$I^{yy} = \frac{3k\mu^2\sin^4\alpha}{8} = \frac{I^{yx}}{3}. \tag{A6}$$

For a randomly oriented sample (0 V on the motor in practice) where we also average over α

$$I^{zy} = I^{zx} = \frac{k\mu^2}{16} \tag{A7}$$

$$I^{yy} = \frac{9k\mu^2}{64} = \frac{I^{yx}}{3}. \tag{A8}$$

Thus we can write the FDFLD on a sample of which a small fraction $\gamma \ll 1$ is oriented as:

$$\begin{aligned} LD_{90} &= \lg\frac{(I^{yy}+I^{yx})}{(I^{zy}+I^{zx})} \\ &= \lg\left(\frac{(1-\gamma)\dfrac{(9+27)}{64}+\gamma\dfrac{(3+9)\sin^4\alpha}{8}}{(1-\gamma)\dfrac{1}{8}+\gamma\sin^2\alpha\cos^2\alpha}\right) \\ &= \lg\left(\frac{(1-\gamma)\dfrac{9}{16}+\gamma\dfrac{3\sin^4\alpha}{2}}{(1-\gamma)\dfrac{1}{8}+\gamma\sin^2\alpha\cos^2\alpha}\right) \\ &\approx \lg\left[\frac{9}{2}\left(1+\gamma\frac{\dfrac{3\sin^4\alpha}{2}-\sin^2\alpha\cos^2\alpha}{(1-\gamma)\dfrac{1}{8}+\gamma\sin^2\alpha\cos^2\alpha}\right)\right] \\ &\approx \lg\left(\frac{9}{2}\right) - \frac{8\gamma\sin^2\alpha}{3\lg 10}(1+2\cos 2\alpha). \end{aligned} \tag{A9}$$

So for transitions polarized along the orientation axis:

$$\alpha = 0 \quad LD_{90} \approx \lg\left(\frac{9}{2}\right), \tag{A10}$$

which after subtraction of the unoriented baseline is zero to first order in γ.

For transitions polarized perpendicular to the orientation axis

$$\alpha = 90^\circ \quad LD_{90} \approx \lg\left(\frac{9}{2}\right) + \frac{8\gamma}{3\lg 10}, \tag{A11}$$

which after subtraction of the unoriented baseline is positive to first order in γ. The sign behavior of the variable term in Eq. A9 that is not removed on baseline subtraction is summarized in Fig. 11.

It is interesting to note that the true sample baseline for all the LD_{90} spectra is not zero. However, if γ is small then this effect is cancelled out when a nonrotating sample signal is subtracted. Another intriguing feature of this experiment's configuration is that because we take the logarithms of a ratio, rather than a simple difference as in, e.g., fluorescence polarization anisotropy, the absolute magnitude of the signals is independent of concentration, although the signal/noise is affected by concentration and fluorescence intensity.

The zero points in the polarization dependent term of LD_{90} are at 0°, 60°, 120°, 180°. Thus one expects the data, as presented in this work with the zero rotation spectrum subtracted, to have a negative signal if the transition moment lies between 0° and 60° from the DNA long axis, and a positive sign if it lies from 60° to 90°. This contrasts with standard configuration LD spectra where the change-over point is the magic angle of 54.7° and the sign dependence is the opposite way around. So intercalators are expected to have a positive LD_{90} signal and long-axis polarized transitions of groove binders to have a negative signal.

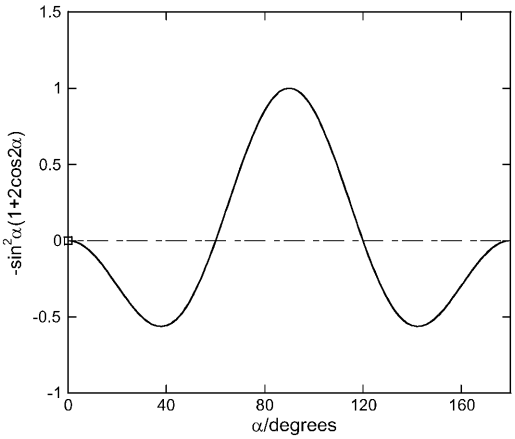


FIGURE 11 Plot of $\sin^2\alpha(1+2\cos 2\alpha)$ from 0° to 180°, although the plot has symmetry ~90°.

The authors gratefully acknowledge the help of the referees as well as discussions with Y. Dupont and S. Windsor in leading to an understanding of the underlying basis of what had been measured. The technical skill and enthusiasm of Rhoderick Mortimore of Crystal Precision Optics in constructing the LD cells is gratefully acknowledged.

Financial support from the Engineering and Physical Sciences Research Council, UK (GR/M91105) and Syngenta, UK was also essential to the progress of this work.

REFERENCES

Bloemendal, M., and R. Vangrondelle. 1993. Linear-dichroism spectroscopy for the study of structural properties of proteins. *Mol. Biol. Rep.* 18:49–69.

Canet, D., K. Doering, C. M. Dobson, and Y. Dupont. 2001. High-sensitivity fluorescence anisotropy detection of protein-folding events: application to α-lactalbumin. *Biophys. J.* 80:1996–2003.

Chou, P., and W. C. Johnson. 1993. Base inclinations in natural and synthetic dyes. *J. Am. Chem. Soc.* 115:1205–1214.

Colson, P., C. Houssier, and C. Bailly. 1995. Use of electric linear dichroism and competition experiments with intercalating drugs to investigate the mode of binding of Hoechst-33258, Berenil and Dapi to Gc sequences. *J. Biomol. Struct. Dyn.* 13:351–366.

Couette, M. 1890. Etudes sur le frottement des liquides. *Ann. Chim. Phys.* 6:433–510.

Dafforn, T. R., J. Rajendra, D. J. Halsall, L. C. Serpell, and A. Rodger. 2004. Protein fibre linear dichroism for structure determination and kinetics in a low-volume, low-wavelength Couette flow cell. *Biophys. J.* 86:404–410.

Eriksson, S. 1992. DNA-Ligand Interactions Studied by Optical Spectroscopy. Chalmers University of Technology and University of Goteburg, Goteborg, Sweden.

Eriksson, S., S. K. Kim, M. Kubista, and B. Nordén. 1993. Binding of 4′,6-diamidino-2-phenylindole (DAPI) to AT regions of DNA: evidence for an allosteric conformational change. *Biochemistry.* 32:2987–2998.

Hofricheter, J., and W. Eaton. 1976. Linear dichroism of biological chromophores. *Annual Review of Biophysical Bioengineering*. 5:511–560.

Houssier, C., B. Hardy, and E. Fredercq. 1974. Interaction of ethidium bromide with DNA. Optical and electrooptical study. *Biopolymers*. 13: 1141–1160.

Johansson, L. B. A., and A. Davidsson. 1985. Analysis and application of linear dichroism on membranes - description of a linear-dichroism spectrometer. *Journal of the Chemical Society-Faraday Transactions I*. 81:1375–1388.

Kim, S. K., S. Eriksson, M. Kubista, and B. Nordén. 1993. Interaction of 4′,6-diamidino-2-phenylindole (DAPI) with poly[d(G-C)$_2$] and poly[d(G-m^{5}C)$_2$]: evidence for major groove binding of a DNA probe. *J. Am. Chem. Soc*. 115:3441–3447.

Kubista, M., B. Akerman, and B. Nordén. 1987. Characterization of interaction between DNA and 4′,6-diamidino-2-phenylindole by optical spectroscopy. *Biochemistry*. 26:4545–4553.

Lee, C., and N. Davidson. 1968. Flow dichroism of deoxyribonucleic acid solutions. *Biopolymers*. 6:531–550.

Lowry, T. M. 1935. Optical Rotatory Power. F. G. Donnan, editor. Longmans, Green and Co., London, UK. 405–408.

Miles, A. J., F. Wien, J. G. Lees, A. Rodger, R. W. Janes, and B. A. Wallace. 2003. Calibration and standardisation of synchrotron radiation circular dichroism and conventional circular dichroism spectrophotometers. *Spectroscopy*. 17:653–661.

Moon, J. H., S. K. Kim, U. Sehlstedt, A. Rodger, and B. Nordén. 1996. DNA structural features responsible for sequence-dependent binding geometries of Hoechst 33258. *Biopolymers*. 38:593–606.

Nordén, B., M. Kubista, and T. Kurucsev. 1992. Linear dichroism spectroscopy of nucleic-acids. *Q. Rev. Biophys*. 25:51–170.

Nordén, B., and T. Kurucsev. 1994. Analysing DNA complexes by circular and linear dichroism. *J. Mol. Recognit*. 7:141–156.

Nordén, B., and F. Tjerneld. 1976. High-sensitivity linear dichroism as a tool for equilibrium analysis in biochemistry. Stability constant of DNA-ethidium bromide complex. *Biophys. Chem*. 4:191–198.

Oriel, P., and J. Schellman. 1966. Studies of the birefringence and birefringence dispersion of polypeptides and proteins. *Biopolymers*. 4:469–494.

Rodger, A. 1993. Linear dichroism. *Methods Enzymol*. 226:232–258.

Rodger, A., J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, et al. 2002. Flow oriented linear dichroism to probe protein orientation in membrane environments. *Physical Chemistry Chemical Physics*. 4:4051–4057.

Simonson, T., and M. Kubista. 1993. DNA orientation in shear flow. *Biopolymers*. 33:1225–1235.

Takakuwa, T., T. Konno, and H. Meguro. 1985. A new standard for calibration of circular dichroism: ammonium *d*-10-camphorsulfonate. *Analytical Sciences*. 1:215–218.

Wada, A. 1964. Chain regularity and flow dichroism of deoxyribonucleic acids in solution. *Biopolymers*. 2:361–380.

Wada, A. 1972. Dichroic spectra of biopolymers oriented by flow. *Applied Spectroscopy Reviews*. 6:1–30.

Wada, A., and S. Kozawa. 1964. Instrument for the studies of differential flow dichroism of polymer solutions. *Journal of Polymer Science Part A*. 2:853–864.

Wells, R. D., J. E. Larson, R. C. Grant, B. E. Shortle, and C. R. Cantor. 1970. Physicochemical studies on polydeoxyribonucleotides containing defined repeating nucleotide sequences. *J. Mol. Biol*. 54:465–497.

Wilson, R., and J. Schellman. 1978. The flow linear dichroism of DNA: comparison with the bead-spring theory. *Biopolymers*. 17: 1235–1248.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Radiation Physics and Chemistry 71 (2004) 1023–1030
Radiation Physics and Chemistry
www.elsevier.com/locate/radphyschem

Ferric and ferrous ions: binding to DNA and influence on radiation-induced processes

Hanna B. Ambroż[a], Terence J. Kemp[b], Alison Rodger[b], Grażyna Przybytniak[a,*]

[a] Institute of Nuclear Chemistry and Technology, Dorodna 16, 03-195 Warsaw, Poland
[b] Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK

Received 1 September 2003; accepted 8 November 2003

Abstract

The role of the two principal oxidation states of iron ions in modifying the response of DNA to ionising radiation depends not only on their contrasting redox character but also on their abilities to induce different conformational changes to the helix. Circular dichroism and linear dichroism studies presented here reveal that Fe(III) induces major stereochemical effects in comparison to Fe(II). This is attributed not only to the charge difference but also to their electronic structure that determines the sites of binding of both iron ions to DNA. It seems that Fe(III), which induces strong changes in the helix geometry, is involved in binding to interior base sites. The influence of Fe(II) is much weaker, thus it is probably preferentially localised at the backbone of DNA strands. Hitherto this has not been unambiguously confirmed experimentally. These stereochemical effects have significant consequences for radiation-induced processes and the electron transfer pathway in this vital biopolymer.

Keywords: DNA radiolysis; Ferric ion; Ferrous ion; EPR; Dichroism

1. Introduction

Iron present in living cells is able not only to fulfil many important biological functions but also to initiate some undesirable processes, e.g. it can contribute to the degradation of biomolecules due to its participation in the Fenton reaction. Consequently, the influence of iron ions in damage to DNA, induced either chemically or by irradiation, has been widely studied (Symons and Gutteridge, 1998; Toyokuni and Sagripanti, 1992; Toyokuni, 1996). The direct effect of ionising radiation creates, in random fragments of DNA, holes and electrons that can move to, and be localised at, other sites. A convenient way to estimate the ratio of electron gain and electron loss centres is to eliminate either the holes or electrons. In the radiation chemistry of bioorganic systems, the iron ions Fe(III) and Fe(II), can play the roles of electron and hole scavengers, respectively. In electron paramagnetic resonance (EPR) experiments carried out at cryogenic temperatures, the paramagnetic ferric ion Fe(III) ($3d^5$ in high spin configuration, $S = 5/2$) appears to be a convenient modifier, as its EPR absorption does not interfere with those of organic radicals, being situated usually at relatively low magnetic field (Symons, 1978). Ferrous ion Fe(II) has an even number of electrons ($3d^6$, high- or low-spin configuration) so under certain conditions does not show EPR absorption.

We have found that Fe(III) ions selectively diminish the population of the radiation-induced, protonated radical anion of thymine and that no electron transfer from the radical anion of cytosine to thymine proceeds in frozen aqueous solution (Ambroż et al., 1998). The electron migration and capture processes are very efficient, because on addition of Fe(III), for a ratio of iron:nucleotide as low as 1:100, the overall

*Corresponding author. Tel.: +48-22-811-23-47; fax: +48-22-811-15-32.
E-mail address: przybyt@ichtj.waw.pl (G. Przybytniak).

0969-806X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.radphyschem.2003.11.003

concentration of paramagnetic species decreases by about 30%. Such a strong effect indicates that Fe(III) must be associated with the DNA helix and the questions as to where and how iron binds to DNA become highly significant. On the other hand, Fe(II) is known as an agent that influences photo- and radiation-induced processes due to its contribution in the production of highly oxidative species (Yurkova et al., 1999; Balazani and Carassiti, 1970). The redox chemistry of iron ions is manifested in aqueous solution in the well-known Fenton reactions. Such solutions are usually treated with a buffer containing ethylene diamine tetraacetatic acid disodium salt (EDTA). Ferrous ions can form complexes with EDTA and one molecule of H_2O, H_2O_2 or O_2, as EDTA is hexadentate and in the ferrous–EDTA complex there is a free co-ordination site (Yurkova et al., 1999; Seibig and van Eldik, 1997), this having been demonstrated crystallographically for the $[Fe(EDTA)(H_2O)]^{2-}$ ion which is a mono-capped trigonal prism (Mizuta et al., 1993). In such a system, the negatively charged Fe(II) species are repelled from the polyanionic biomolecules, e.g. DNA, and if they are involved in generation of $OH^{\bullet}$, the non-specific damage of solute appears. However, in those experiments where Fe(II) complexes are positively charged, there is a high probability that the iron is bound directly to DNA and site-specific damage might be expected. Hence the localisation of the iron ions can play a crucial role in the outcomes of some Fenton-like reactions.

The character of binding of iron ions with nucleotides and nucleic acids has remained unclear. The influence of Fe(III) on adenine nucleotide and its derivatives has been studied by Mössbauer spectroscopy (Rabinowitz et al., 1966). The results support the view that ferric ions can be bound at purine nitrogen atoms at neutral and basic pH but not under acidic conditions since Fe(III) is then coordinated octahedrally by oxygen atoms. This would imply binding of iron to the DNA backbone, very probably at the phosphate groups. Equilibrium constant data for the interaction of ferric iron with adenosine and other nucleoside di- or tri-phosphates showed that there is little or no interaction of Fe(III) with the bases (Goucher and Taylor, 1964). Spin-lattice relaxation time studies also led to the conclusion that iron(III) ions bond to phosphate groups rather than to bases (Eisinger et al., 1962). Contrary to these suggestions, Signer (1964) reported that Fe(III) binds to the bases. Information on metal ion–nucleic acid interactions conducted with different techniques has been collected in two reviews published in the 1970s (Sissoëff et al., 1976; Izatt et al., 1971) and the papers cited therein reveal discrepancies between authors concerning Fe–nucleic acid binding. Studies on the interaction of metal ions with synthetic DNA indicated that Fe(III), like some other transition and other metal ions, Mn(II), Cu(II), Cd(II), In(III) and Pb(II), induces monophasic structural transition in double-stranded poly d(G - - C) (Rossetto and Nieboer, 1994). A recent monograph on nucleic acid structure, properties and functions, although discussing in detail metal ion binding to nucleic acids, does not quote any information on the effect of iron ions on the stability and conformation of the double helix (Bloomfield et al., 2000).

To understand the influence of ferric and ferrous ions on DNA damage induced by ionising radiation it is necessary to determine their localisation, especially as regards proximity to the DNA helix. Circular dichroism (CD) and linear dichroism (LD) are very informative in this respect as such measurements can provide data concerning the binding of metal ions in relation to chiral biomolecules, and have been employed successfully to investigate the stability of particular DNA conformations. Our experiments were performed at various concentrations of iron ions, both in buffered and unbuffered DNA solutions, and for systems both containing and without EDTA.

2. Materials and methods

Calf thymus DNA used in this study for EPR, CD and LD experiments is commercially available (Sigma) and was used without further purification. Ferric and ferrous chlorides were purchased from Avocado Research Chemicals Ltd. Other chemicals: EDTA; Tris (tris(hydroxymethyl) aminomethane)–EDTA buffer; sodium chloride; and cacodylic acid sodium salt were obtained from Sigma. DNA was dissolved in water and refrigerated (4°C). As a freshly prepared solution of $FeCl_3$ caused DNA precipitation even at low concentrations, the reagent was stored a few hours before use to achieve equilibrium with all potential ligands. The samples were prepared 24 h before experiments by mixing solutions of DNA and equilibrated $FeCl_3$ in appropriate proportions. The $FeCl_2$ solutions were freshly prepared and added to DNA directly before measurements to avoid oxidation of Fe(II). Cacodylate buffer, pH = 6.8, was made up from $C_2H_6AsO_2Na$ and hydrochloric acid according to Gomori (1955) and subsequently dissolved in NaCl solutions. Ultrapure water (18.2 MΩ) was used in all experiments. The flow cuvette cell is detailed by Rodger (1993). The concentration of DNA was expressed in terms of molarities of nucleotides and was determined using $\varepsilon_{257} = 6600\,mol^{-1}\,dm^3\,cm^{-1}$.

Samples for EPR experiments were prepared from air-saturated aqueous solutions frozen in the form of rods in liquid nitrogen and irradiated at 77 K with a ^{60}Co source to a dose of 10 kGy at a dose rate of 2.3 kGy/h. After irradiation, EPR spectra were measured with a Bruker X-band ESR 300 spectrometer using a microwave power of 15 μW. EPR spectra were recorded at 77 K prior to,

and following brief warming above 77 K, to the required temperature controlled by a thermocouple. APOLLO software was applied to analysis of experimental spectra (Dusemund, 1998). The procedure consists of double integration of recorded EPR spectra for the determination of the decrease of relative concentration of radicals versus annealing temperature.

CD and LD spectra were recorded using a Jasco J-175 spectropolarimeter adapted for LD measurements.

Plasmid DNA pBR322 used for electrophoresis was supplied by MBI Fermentas (buffered with 10 mM Tris-HCl, pH = 7.6, and 1 mM EDTA). Samples of DNA were prepared 24 h before experiments, kept at 4°C and then irradiated in the absence and presence of iron(III) ions at room temperature with the Mineyola source at a dose rate of 37.3 Gy/h to the doses indicated below. The irradiated samples consisted of the following components: 0.6 mM nucleotides, 0.4 mM EDTA, 4 mM Tris and $FeCl_3$ in the range of concentration 0–1 mM. Before electrophoresis the samples were mixed with gel loading buffer which contained Ficoll, bromophenol blue and xylene cyanol FF. The gel was prepared from agarose (Serva) and Tris-acetate-EDTA (TAE) buffer, pH = 8.0. Ethidium bromide was added to the agarose solution to a final concentration of 0.5 μg/ml. Electrophoresis was carried out for 2 h under TAE running buffer at a voltage of 100 V. The results were recorded with a digital camera, saved in a computer and examined by applying DNAnalysis software (Mirkowski, 1998). Upon determination of the percentage of supercoiled (undamaged), circular (single strand break (ssb)) and linear (double strand break (dsb)) forms, the fraction of supercoiled plasmid was increased 1.5 times as undamaged plasmid binds $\frac{1}{3}$ less ethidium bromide than the other two forms (Spotheim-Maurizot et al., 1992). Experiments were repeated three times and average results utilised. Details of the electrophoresis procedure have been described previously (Ambroż et al., 2001).

3. Results

3.1. Influence of iron(III) ions on DNA radiolysis

The effect of ionising radiation on DNA in the presence of Fe(III) was probed both by EPR spectroscopy which detects the radicals, and by gel electrophoresis measurements of DNA single and double strand cleavage. The low-temperature EPR experiments give an indication of the transient species produced by irradiation, whereas the cleavage experiments probe the end effects. Induced by radiolysis, primary radicals convert to the secondary products which are eventually responsible for losing of DNA integrity. Therefore both applied methods deliver complementary information on the effects induced by ionising radiation.

In EPR experiments frozen aqueous solutions of only two components, DNA and iron(III) ions, were examined. Under such conditions the changes observed on radiolysis of DNA result only from the influence of Fe(III). Samples of DNA and DNA doped with $FeCl_3$ were irradiated at 77 K and annealed to 150 K (to eliminate $OH^{\bullet}$ absorption). Under these conditions the EPR spectra (Fig. 1A and B) show the direct effect of ionising radiation on DNA. At 150 K in the presence of Fe(III) there is a reduction in the doublets attributed to the radical anions of thymine and cytosine. The decrease in the radical anion concentrations results in elimination

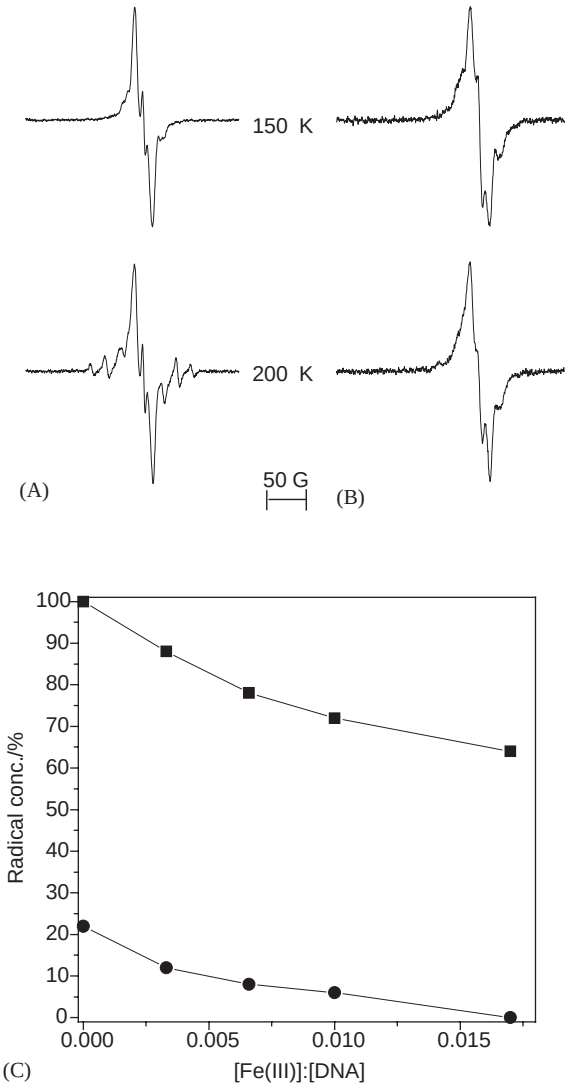


Fig. 1. EPR spectra of frozen aqueous solutions of DNA irradiated to a dose of 10 kGy (A) without additives, and (B) in presence of 5 mM Fe(III) ([Fe(III)]: [DNA] = 1:60). (C) concentration of total DNA radicals (■) and $TH^{\bullet}$ (●) as a function of ratio iron(III) ions per nucleotide.

of secondary species formed by protonation of the radical anion of thymine ($T^{\bullet -}$), namely the thymyl radical ($TH^{\bullet}$), which exhibits a highly characteristic octet. Therefore, as seen in Fig. 1B upon annealing to higher temperatures, this product does not appear in the presence of 5 mM Fe(III) which shows that electrons generated in DNA following irradiation are effectively scavenged preferentially by iron ions rather than by thymine. Quantitative analysis of the results indicates that Fe(III) reduces the yield of DNA radicals significantly. A gradual decrease of the concentrations of total DNA radicals and $TH^{\bullet}$ versus iron(III) content on annealing to 200 K is depicted in the graph in Fig. 1C. At the ratio [Fe(III)]:[DNA] = 1:60, the generation of all paramagnetic intermediates is reduced by ca. 35%. This phenomenon must result not only from efficient electron transfer through the DNA π-system but also from the convenient localisation of ferric ion which facilitated the process. It seems that ferric ions do not influence the formation of DNA products generated via the oxidative pathway, i.e. the radical cation of guanine which exhibits a singlet, and changes can be observed only as regards the production of radical anions.

The application of electrophoresis for the estimation of DNA cleavage in the presence of Fe(III) brings specific problems. In this case, apart from iron ions and DNA, there are in the system many other components resulting from the methodology. Plasmid DNA necessary for electrophoresis is delivered in EDTA–Tris buffer. In order to avoid exchange of the buffer leading to the inevitable weight-loss of plasmid and to undesirable breaks of the DNA strands, some studies have been carried out using pristine solutions. In such a medium every investigation which requires an admixture of iron ions involves a contribution of the buffer components to the observed processes. Specially important seems to be the role of EDTA due to its strong complexing ability towards both forms of iron ions.

The effects of [Fe(III)] on ssb and dsb of plasmid DNA in EDTA-containing buffer, following radiolysis at room temperature and separation by electrophoresis of supercoiled, nicked and linear forms of the plasmid, are shown in Fig. 2. The concentration of EDTA is constant at 0.4 mM. The results are presented as ssb and dsb versus Fe(III) concentration at various radiation doses. The extent of the breaks grows with increasing concentration of Fe(III) and with absorbed dose. Up to a dose of 100 Gy, the increase in ssbs is linear with dose, while above that dose the character of the correlation of ssb versus [Fe(III)] becomes inversely proportional. This is presumably due to the formation of two ssbs sufficiently close on both strands to generate a dsb. Below a dose of 50 Gy no dsb was observed. Their prevalence grows much more slowly than ssb, and remains linear over the whole concentration range of iron(III) ions examined.

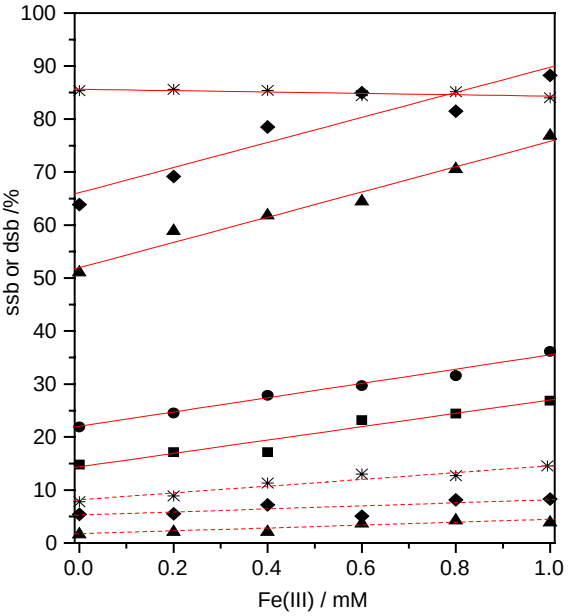
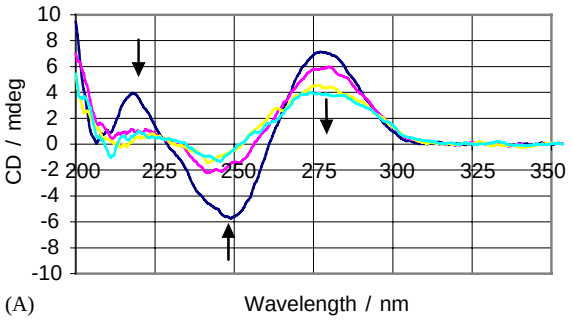


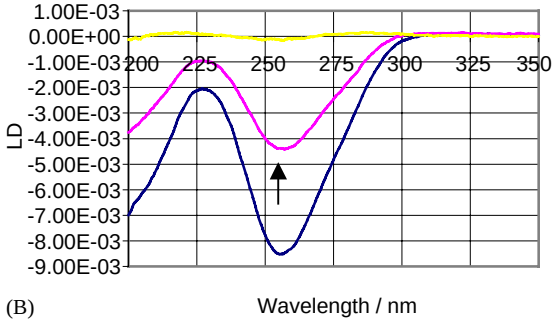
Fig. 2. The influence of Fe(III) on DNA strand breaks induced by ionising radiation at room temperature. (■) 0 Gy, (●) 10 Gy, (▲) 50 Gy, (♦) 100 Gy, (✳) 200 Gy. Full line—ssb, broken line—dsb. The samples were buffered with EDTA-Tris buffer, pH = 7.6.

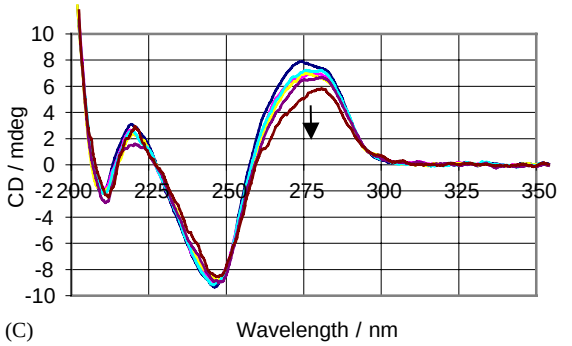
3.2. Circular and linear dichroism of DNA in the presence of Fe(III)

In order to try to understand how the Fe(III) was affecting the DNA, CD and LD experiments were undertaken. Calf thymus DNA in distilled water exhibits CD and LD spectra characteristic for right handed B-DNA (Fig. 3) (Rodger and Nordén, 1997; Rodger et al., 2000). The CD signal (Fig. 3A) consists of two positive bands with λ_{max} at 277 and 217 nm together with a negative band centred at 247 nm. The spectra below 200 nm were not recorded due to the strong Fe(III) absorption interfering with the DNA signal in this region. Increasing concentrations of $FeCl_3$ induce conformational changes in DNA as shown by significant progressive reduction of the signals, as well as a wavelength shift of the positive band. The highest concentration of $FeCl_3$ used in the experiment was [Fe(III)]:[DNA] equal to 3:4 as larger amounts of ferric ions cause precipitation of DNA as an amorphous yellow precipitate. We have not observed condensation of the DNA as detected by UV–Visible absorbance spectroscopy.

LD spectra show the differences in absorption of light that is linearly polarised parallel and perpendicular to the orientation axis. For flow oriented DNA, where the DNA helix is the orientation axis and the transitions are polarised in the plane of the DNA bases, one expects a large negative signal at the absorbance maximum. This







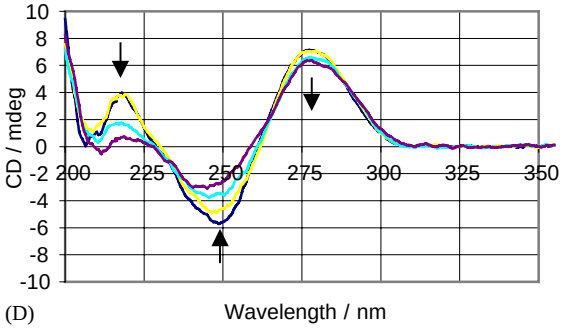


Fig. 3. Spectra of solutions containing 100 μM DNA. Panel A: CD spectra in presence of 0, 25, 50 and 75 μM $FeCl_3$. Panel B: LD spectra in presence of 0, 25 and 50 μM $FeCl_3$. Panel C: CD spectra in 1 mM cacodylate buffer, pH = 6.8, and 5 mM NaCl in presence of various concentration of $FeCl_3$–0, 25, 50, 75, 100 and 200 μM. Panel D: CD spectra of 100 μM DNA at pH = 6.9; 6.1; 5.7; 5.0 adjusted by hydrochloric acid.

signal can be used to monitor the orientation of the DNA helix. As shown in Fig. 3B the orientation of the helix is considerably modified by ferric chloride, and at a $FeCl_3$ concentration as low as half that of [DNA] the signal is completely lost. This means that the LD signals are more sensitive to the presence of iron(III) than the CD spectra. Although a reduction in LD was expected upon addition of the highly charged Fe(III) (Rodger et al., 2000), the effect observed was more dramatic than previously observed.

Since, the addition of $FeCl_3$ up to 75 μM lowers the pH of solutions from 6.8 to ca. 5.0, it was necessary to probe the role of pH in this process. The spectra of DNA: (i) with various additions of ferric chloride in buffered solution at pH 6.8 and increased ionic strength by the addition of 5 mM NaCl; and (ii) without iron on gradual adjusting of the pH up to 5.0 by titration with HCl, are presented in Fig. 3C and D, respectively. Some authors claim that the stability of the helix is relatively unchanged between pH 5 and 9 (Bloomfield et al., 2000). However, in our system the changes in the bands at 247 and 217 nm with varying pH are distinct. Comparison of the results presented in panels A and D shows that the significant reduction of the maximum at 277 nm, together with the decrease of the negative band at 247 nm, results mainly from the influence of iron(III) ions on DNA. Thus we may conclude that although pH does play a role in the effects observed, the effect of Fe(III) is much the most significant.

In unbuffered solutions doped with EDTA the influence of iron(III) ions on DNA conformation depends on the relative concentrations of the components. In the presence of EDTA at high concentration ([EDTA] > [DNA]) (Fig. 4A), the CD intensities are reduced and the shape of the spectrum is different. The first positive band at ca. 280 nm shifts to 285 nm and the second smaller minimum at ca. 220 nm is slightly lower. Moreover, a new positive band emerges around 260 nm. The CD spectra of DNA in solution with 80 μM Fe(III) at various concentrations of EDTA are collected in Fig. 4B. If the concentrations of EDTA and Fe(III) are the same then the character of the CD spectrum remains similar to that which was recorded for the system without admixture of the complexing agent. The excess of EDTA in the presence of the Fe(III) induces dramatic variations in DNA chirality as seen from the different shape of the spectrum for the ratio [EDTA]: [Fe(III)] = 3:1. However the same experiment repeated for the buffered solutions, pH ≅ 7, confirmed that under these conditions the Fe(III)/EDTA complex does not perturb DNA chirality.

In Fig. 4C the differences in the LD spectra are depicted as a function of [Fe(III)] at an EDTA concentration twice that of [DNA]. The signal preserves the same character as in the system without EDTA although the effect observed is weaker.

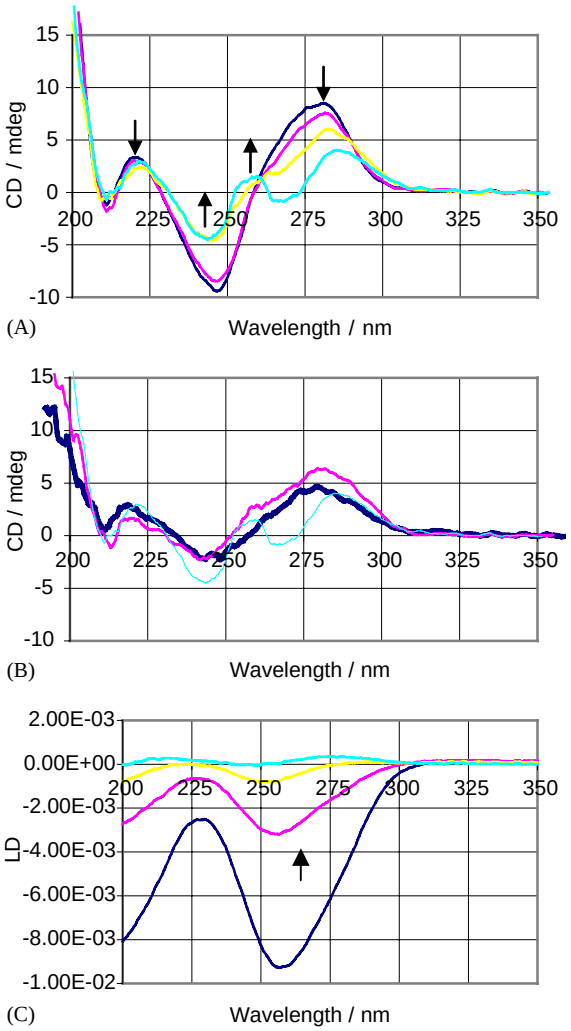


Fig. 4. Panel A: CD spectra of 130 μM DNA solution in 320 μM EDTA, in presence of 0, 20, 40 and 80 μM $FeCl_3$. Panel B: CD spectra of DNA containing 80 μM $FeCl_3$ and 0 μM EDTA (thick line), 80 μM EDTA (standard line) and 320 μM EDTA (thin line). Panel C: LD spectra of 100 μM DNA in 200 μM EDTA in presence of 0, 25, 50 and 75 μM $FeCl_3$.

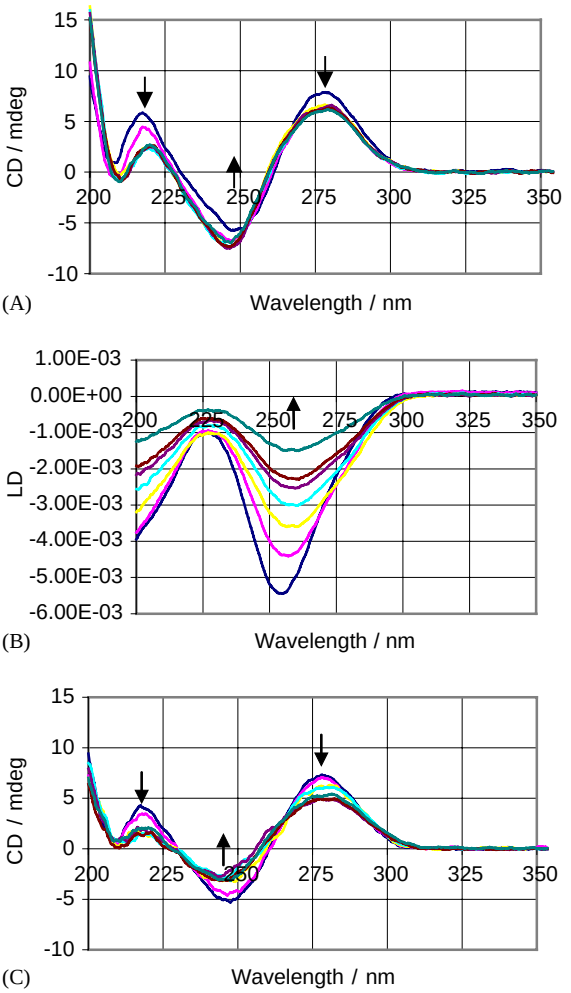


Fig. 5. Spectra of solutions containing 100 μM DNA. Panel A: CD spectra in presence of 0, 50, 100, 200, 400 μM $FeCl_2$. Panel B: LD spectra in presence of 0, 25, 50, 75, 100, 200 and 400 μM $FeCl_2$. Panel C: as in Panel A in presence of 50 μM EDTA, unbuffered solutions.

3.3. *Circular and linear dichroism of DNA in the presence of Fe(II)*

During radiolysis, ferric ions, effectively scavenging electrons, undergo reduction to ferrous species. Therefore it was important to investigate also the binding of Fe(II) to DNA. The CD spectra reveal that the interaction of $FeCl_2$ with DNA is weaker than with Fe(III) (Fig. 5A). The changes in conformation are negligible even at concentrations of ferrous ions much higher than possible with ferric ions. Fig. 5B shows LD signals of DNA on addition of $FeCl_2$. The gradual loss of the signal reflects progressive reduction in DNA orientation in accord with the literature (Rodger, 1993), despite the lower pH (at 400 μM $FeCl_2$ pH decreases to 5.6). For aerated samples containing EDTA (Fig. 5C), the Fe(II)/EDTA complex is oxidised to Fe(III)/EDTA in the course of seconds (Lambeth et al., 1982). The intensities of all peaks are lower with increasing concentration of ferrous ions but less so than observed in the case when Fe(III) is directly added to the system.

4. Discussion and conclusions

The aim of the work reported in this paper has been to explore the consequences of Fe(III) localisation on damage to DNA induced by ionising radiation.

CD/LD spectroscopic experiments were used to probe the conformation of the DNA and the complexation of the iron with DNA. A key issue is the nature of the Fe(III) species in aqueous solution. Iron(III) chloride solutions are acidic due to hydrolysis which prompts the drop in pH according to the equilibria (Cotton and Wilkinson, 1980):

$$[Fe(H_2O)_6]^{3+}+H_2O = [Fe(OH^-)(H_2O)_5]^{2+}+H_3O^+, \quad K = 10^{-3.05},$$

$$[Fe(OH^-)(H_2O)_5]^{2+}+H_2O = [Fe(OH^-)_2(H_2O)_4]^{+}+H_3O^+, \quad K = 10^{-6.31}.$$

The acidity cannot be the main reason for the changes observed in the CD and LD spectra, as: (i) the alterations in CD for DNA solutions in a moderately acidic environment (pH > 5) are much smaller than that observed upon addition of $FeCl_3$ (Figs. 3A and D); (ii) the similar reduction in pH induced by addition of $FeCl_2$ (Fig. 5A) results in negligible modification of the CD signals. Variations in ionic strength of the solution also are not responsible for these effects because the influence of ferric ions on DNA is similar in aqueous solutions at up to 20 mM NaCl (results not shown). Since effects of these two factors are not the main influences on the DNA chirality then, as CD and LD showed, the cationic Fe(III) species play a crucial role in these effects.

Release of a proton from the Fe(III) complex results in the appearance of OH^- in the inner coordination shell of iron which facilitates the exchange of water molecules (Biruš et al., 1993). The proportions of the major species of Fe(III) hydrolysis depend on pH as ligands can undergo exchange with Cl^- until equilibrium is established including species such as $[Fe(Cl^-)(H_2O)_5]^{2+}$ and $[Fe(OH^-)(Cl^-)(H_2O)_4]^{+}$. In neutral and basic solutions these reaction equilibria are shifted to the right and eventually lead to the appearance of $[Fe(OH^-)_3(H_2O)_3]$ or its Cl^--substituted derivatives, and their precipitation (Nicholls, 1973). It has been shown that the less protonated (or more negative) the ligands, the faster is the release of water molecules from the coordination sphere of the ferric ion and the higher its reactivity (Biruš et al., 1993; Nicholls, 1973). Consequently, for such species, replacement of water molecules by functional groups of DNA is expected.

From EPR studies it is known that Fe(III) is able to scavenge electrons induced in DNA, even for a very small ratio [Fe(III)]:[DNA]. We have shown that one Fe(III) species per 3 turns of the DNA helix (assuming all are DNA bound) quenches about 70% of the radicals formed via the reductive pathway, suggesting that the effect of a bound 'quencher' spans more than one turn of the helix. However, in frozen aqueous solution even very high levels of iron ions are unable to eliminate all transients produced in the reduction of DNA (Lange et al., 1995). This indicate that electron transfer through the π–system of the bases is limited. On the basis of our CD and LD studies, we may state that the larger the ferric ion concentration, the greater is effect induced on the DNA double-helical structure. Thus, we conclude that at higher concentration of Fe(III) inhibition of electron transfer results from the breakdown of the DNA helix structure. On the other hand, the significant role of Fe(III) in the modification of the DNA chirality and radiolysis indicates that the ions may be coordinated to the DNA bases.

In contrast to ferric ions, the effect of Fe(II) on the DNA is almost non-existant. In neutral solutions, $[Fe(H_2O)_6]^{2+}$ barely undergoes deprotonation as the pK of its first stage is probably above 7 (there are major discrepancies between various authors, who mostly place the pK value in the region 7–9.5 (Stuglik, 1978)). Fe(II) ions coordinated by neutral molecules barely dissociate so there is limited possibility for covalent binding to the DNA bases. It is most likely therefore that the Fe(II) species interact with the DNA phosphate groups.

The logarithm of the binding constant of Fe(III) to EDTA is 25.1 (Tilbrook and Hider, 1998) and the addition of EDTA is expected to remove all forms of Fe(III) from the vicinity of the DNA as the resulting ion at neutral pH, $[Fe(EDTA)H_2O]^-$, is negatively charged. During radiolysis such a complex can be reduced by electrons generated, and subsequently increase the level of strand breaks, producing highly oxidising species. Even when the concentration of Fe(III) exceeds that of EDTA no discontinuity is observed in ssb and dsb changes (Fig. 2). Therefore we suggest that when the Fe(II) species are formed due to radiolytic reduction of Fe(III), both chelated and non-chelated by EDTA, Fenton-like reactions proceed with similar yields.

At lower pH, the CD spectra clearly show interaction of Fe(III)/EDTA complexes with DNA. The new band emerging around 260 nm (Fig. 4B) confirms the interaction between doped components and DNA in acidic solution. The absorption becomes more evident for higher concentrations of EDTA (and growing acidity of the solution), in spite of neligeable consequences of EDTA presence on the DNA structure if ferric ions are absent in the system. This presumably correlates with the presence of a DNA-binding species produced upon protonation of the Fe(III) ligands. Damage to the DNA helix taking place while the solution is not kept neutral by buffering seems to imply that the gradual protonation of EDTA bound to Fe(III) eventually may lead to products which form a new chiral centre with the DNA.

Acknowledgements

We thank the Royal Society for facilitating this work under its Joint Project Grant Scheme for Central and

Eastern Europe. We also thank Professor J. Narbutt for his helpful comments.

References

Ambroż, H.B., Kemp, T.J., Kornacka, E.M., Przybytniak, G.K., 1998. The role of copper and iron ions in the γ-radiolysis of DNA. Radiat. Phys. Chem. 53, 491–499.

Ambroż, H.B., Bradshaw, T.K., Kemp, T.J., Kornacka, E.M., Przybytniak, G.K., 2001. Role of iron ions in damage to DNA: influence of ionising radiation, UV light and H_2O_2. J. Photochem. Photobiol. A 142, 9–18.

Balazani, V., Carassiti, V., 1970. Photochemistry of Coordination Compounds. Academic Press, London, New York.

Biruš, M., Kujundžić, N., Pribanić, M., 1993. Kinetics of complexation of iron (III) in aqueous solution. Prog. React. Kinet. 18, 171–271.

Bloomfield, V.A., Crothers, D.M., Tinoco, I., 2000. Nucleic acids. Structures, Properties, and Functions. University Science Book, Sausalito, CA.

Cotton, F.A., Wilkinson, G.W., 1980. Advanced Inorganic Chemistry. Wiley, New York.

Dusemund, B., 1998. Ph.D. Thesis, University of Saarland.

Eisinger, J., Shulman, R.G., Szymanski, B.M., 1962. Transition metal binding in DNA solutions. J. Phys. Chem. 36, 1721–1728.

Gomori, H., 1955. Preparation of buffers for use in enzyme studies. Methods Enzymol. 1, 138–146.

Goucher, C.R., Taylor, J.F., 1964. Compounds of ferric iron with adenosine triphosphate and other nucleoside phosphates. J. Biol. Chem. 239, 2251–2255.

Izatt, R.M., Christensen, J.J., Rytting, J.H., 1971. Sites and thermodynamic quantities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides. Chem. Rev. 71 (5), 439–447.

Lambeth, D.O., Ericson, G.R., Yorek, M.A., Ray, P.D., 1982. Implications for in vitro studies of the autoxidation of ferrous ion and the iron-catatyzed autoxidation of dithiothreitol. Biochim. Biophys. Acta 719, 501–508.

Lange, M., Weiland, B., Hüttermann, J., 1995. Influence of electron scavengers on the radical formation in thymidine-5′-monophosphate and DNA in frozen aqueous solution and glasses. Int. J. Radiat. Biol. 68 (4), 475–486.

Mirkowski, J., 1998. Program do analizy zdjec elektroforezy DNAnalyser. Raport IChTJ, B 4/98 (in Polish).

Mizuta, T., Wang, J., Miyoshi, K., 1993. A seven-coordinate structure of Iron(II)-ethylenediamine-N,N,N^1,N^1-tetraacetato complex as determined by X-ray crystal analysis. Bull. Chem. Soc. Jpn. 66, 2547–2551.

Nicholls, D., 1973. Compounds of iron. In: Bailar, J.C., Emeleus, H.J., Nyholm, R., Trotman-Dickenson, A.F. (Eds.), Comprehensive Inorganic Chemistry, Vol. 3. Pergamon Press, Oxford.

Rabinowitz, I.N., Davis, F.F., Herber, R.H., 1966. Mössbauer-effect studies on metal binding in purine compounds. J. Am. Chem. Soc. 88, 4346–4354.

Rodger, A., 1993. Linear dichroism. Methods Enzymol. 226, 232–258.

Rodger, A., Nordén, B., 1997. Circular Dichroism and Linear Dichroism. Oxford University Press, Oxford.

Rodger, A., Sanders, K.J., Hannon, M.J., Meistermann, I., Parkinson, A., Vidler, D.S., Haworth, I.S., 2000. DNA structure control by polycationic species: polyamine, cobalt amine and di-metallo transition metal chelates. Chirality 12, 221–236.

Rossetto, F.E., Nieboer, E., 1994. The interaction of metal ions with synthetic DNA: induction of conformational and structural transitions. J. Inorg. Biochem. 54, 167–186.

Seibig, S., van Eldik, R., 1997. Kinetics of [Fe^{II}(EDTA)] oxidation by molecular oxygen revisited. New evidence for a multistep mechanism. Inorg. Chem. 36, 4115–4120.

Signer, B., 1964. Studies of complex of ribonucleic acid from tobacco mosaic virus with Ag^+, Hg^{2+}, Ca^{2+} and Fe^{3+}. Biochim. Biophys. Acta 80, 137–144.

Sissoëff, I., Grisvard, J., Guillé, E., 1976. Studies on metal ions-DNA interactions: specific behaviour of reiterative DNA sequences. Prog. Biophys. Mol. Biol. 31, 165–199.

Spotheim-Maurizot, M., Garnier, F., Sabatier, R., Charlier, M., 1992. Metal ions protect DNA against strand breakage induced by fast neutrons. Int. J. Radiat. Biol. 62, 659–666.

Stuglik, Z., 1978. Ph.D. Thesis, Institute of Nuclear Research, Warsaw.

Symons, M.C.R., 1978. Chemical and Biochemical Aspects of Electron Spin Resonance Spectroscopy. Van Nostrand Reinhold Co, New York.

Symons, M.C.R., Gutteridge, J.M.C., 1998. Free Radicals and Iron: Chemistry, Biology and Medicine. Oxford Science Publications, Oxford University Press, Oxford.

Tilbrook, H., Hider, L., 1998. Iron chelators for clinical use. In: Sigel, A., Sigel, H. (Eds.), Metal Ions in Biological Systems. Iron Transport and Storage in Microorganisms, Plants and Animals, Vol. 35. Marcel Dekker, New York.

Toyokuni, S., Sagripanti, J.-L., 1992. Iron-mediated DNA damage: sensitive detection of DNA strand breakage catalyzed by iron. J. Inorg. Biochem. 47, 241–248.

Toyokuni, S., 1996. Iron-induced carcinogenesis: the role of redox regulation. Free Radical Biol. Med. 20 (4), 553–566.

Yurkova, I.L., Schuchmann, H.-P., von Sonntag, C., 1999. Production of OH radicals in the autoxidation of the Fe(II)–EDTA system. J. Chem. Soc., Perkin Trans. 2, 2049–2052.

J|A|C|S
ARTICLES
Published on Web 08/20/2004

Flow Linear Dichroism to Probe Binding of Aromatic Molecules and DNA to Single-Walled Carbon Nanotubes

Jascindra Rajendra,[†] Mark Baxendale,[‡] Laurence Georges Dit Rap,[†] and Alison Rodger*[,†]

Contribution from the Department of Chemistry, University of Warwick, Coventry, CV4 7AL, U.K., and Department of Physics, Molecular and Materials Physics Group, Queen Mary, University of London, Mile End Road, London E1 4NS, U.K.

Received March 5, 2004; E-mail: a.rodger@warwick.ac.uk

Abstract: Structures of carbon nanotube/ligand complexes were studied by flow linear dichroism (the differential absorption of light polarized parallel and perpendicular to the flow orientation direction) with the aim of establishing linear dichroism as a technique to study such systems. Anthracene, naphthalene, and DNA were chosen as ligands, and the potential for flow linear dichroism to probe ligands noncovalently (as well as covalently) bound to single-walled nanotubes is reported. Linear dichroism enables the determination of approximate orientations of the ligands on the carbon nanotubes.

1. Introduction

Carbon nanotubes (CNTs), as a new allotrope of carbon with extraordinary properties and many captivating potential applications, have attracted much interest since they were discovered in 1991 by Iijima.[1] Carbon nanotubes are fascinating structures for fundamental science and technological applications in the new field of nanoscience.[2] Due to their unusual geometry and their structural and electronic properties, these carbon nanostructures are viewed as promising building blocks for molecular electronics and as reinforcing fillers in composite materials. A lot of research targeting applications of multiwalled (MW) and single-walled (SW) CNTs (SWNTs) has been undertaken, including their utilization as field emitters, electrodes of lithium-ion batteries, tips for scanning probe microscopes (SPM), nanoelectronic devices, and supports in catalysis.[3] Recently, significant progress has been made in both the chemical functionalization and manipulation of CNTs.[4] In this work we have been focusing on the noncovalent functionalization of CNTs.

Since the discovery of carbon nanotubes, thousands of papers have been published on MW and SW CNTs but the field is in desperate need of new structural characterization techniques for CNT systems. Previous work has quantified the structural anisotropy that develops in surfactant-stabilized aqueous nanotube suspensions under steady shear flow by the combination of small-angle polarized light scattering and optical microscopy in situ.[5,6] Chemisorptive attachment of anthracene and several derivatives onto the sidewalls of SWNTs has been investigated using FT-IR, fluorescence, and absorbance spectroscopies, and some discussion of the changes in spectra of SWNT−anthracene absorptive adducts compared to free anthracene and its derivatives has been given.[7] Absorbance data probing the effect of anthracene adsorption on SWNTs has been determined between 300 and 400 nm, where anthracene has pronounced vibronic structure in its electronic absorbance spectrum. In a related application, the physisorption of aromatic molecules on the graphite basal plane has been investigated by NMR experiments, neutron diffraction, and Penning ionization electron spectroscopy.[8,9] However, structural information about physisorbed films of naphthalene and anthracene molecules on graphite is still lacking, and these approaches have yet to be extended to CNTs. In this paper we report the first use of flow linear dichroism with SWNTs and SWNT−ligand adducts and show it has potential for probing ligands noncovalently (as well as covalently) bound to the SWNT. With the use of UV−vis flow linear dichroism on an adapted Jasco J-715 circular dichroism spectropolarimeter we are able to probe the interaction over a wide wavelength range (190−800 nm) and thus use polarized spectroscopy to give more information about the nature of the ligand−CNT interactions.

Interpretation of the flow LD data ideally requires the transition moments of the species being studied to be known. In the case of nanotubes, there is very little information about their spectroscopy in the near and far UV regions, which are

* Corresponding author. Phone: -44-24-76523234. Fax: -44-24-76524112.
[†] University of Warwick.
[‡] University of London.

(1) Liang, Q.; Gao, L. Z.; Li, Q.; Tang, S. H.; Liu, B. C.; Yu, Z. L. *Carbon* **2001**, *39*, 897−903.
(2) Bernier, P.; Maser, W.; Jouret, C.; Loiseau, A.; de la Chapelle, M. L.; Lefrant, S.; Lee, R.; Fischer, J. E. *Carbon* **1998**, *36*, 675−680.
(3) Saito, T.; Matsushige, K.; Tanaka, K. *Physica B: Condens. Matter* **2002**, *323*, 280−283.
(4) Azamian, B. R.; Davis, J. J.; Coleman, K. S.; Bagshaw, C. B.; Green, M. L. H. *J. Am. Chem. Soc.* **2002**, *124*, 12664−12665.
(5) Hobbie, E. K.; Wang, H.; Kim, H.; Han, C. C.; Grulke, E. A.; Obrzut, J. *Rev. Sci. Instrum.* **2003**, *74*, 1244−1250.
(6) Hobbie, E. K.; Wang, H.; Kim, H.; Lin-Gibson, S.; Grulke, E. A. *Phys. Fluids* **2003**, *15*, 1196−1202.
(7) Zhang, J.; Lee, J. K.; Wu, Y.; Murray, R. W. *Nano Lett.* **2003**, *3*, 403−407.
(8) Tabony, J.; White, J. W.; Delachaume, J. C.; Coulon, M. *Surf. Sci. Lett.* **1980**, *95*, L282−L288.
(9) Bondi, C.; Baglioni, P.; Taddei, G. *Chem. Phys.* **1985**, *96*, 277−285.

10.1021/ja048720j CCC: $27.50

the regions of particular interest for the ligands we have chosen to study. A few papers have been published on transition moment polarizations of fullerenes determined using induced circular dichroism[10] or on optical measurements with polarized light.[11] In the case of CNTs, it generally seems to be assumed that there is no UV spectroscopy and the focus has been on near infrared, infrared, and Raman work. In fact, as discussed below, the CNTs do have UV absorbance; it is just very unstructured and often masked by light scattering.

In addition to aromatic molecules, e.g., naphthalene and anthracene, as nanotube ligands, we were interested in probing the binding of DNA to CNTs. There are reports in the literature of strong interactions between DNA and carbon nanotubes;[12] however, the nature of such interactions is not yet clear. Techniques used to study DNA/nanotube systems have included UV and IR absorbance spectroscopy[13,14] and surface-enhanced infrared absorption (SEIRA)[15] as well as transmission electron microscopy (TEM). None of these techniques give information on how the DNA molecules interact with CNTs, and in most cases the CNT/DNA system has been deposited on a solid surface. A recent report of DNA solubilizing nanotubes uses AFM and absorbance spectroscopy to prove this is in fact the case; however, there was no attempt to characterize the interaction.[16] One motivation for studying the interaction of CNTs with DNA is the potential use of CNTs as atomic force microscopy tips to probe biomacromolecules.[17] In this context or independently, any means of immobilizing DNA onto solid supports may provide a means of using DNA as a recognition agent, e.g., for complementary DNA or for specific proteins.

2. Materials and Methods

Materials. SWNTs were obtained from Dynamic Enterprises Ltd. These had been synthesized by the catalytic arc discharge method.[18] To overcome problems due to inhomogenity of the samples, stock solutions of SWNT (0.5 mg/mL) were prepared. The stock solution of SWNTs was obtained by sonicating the SWNTs for 2 min in aqueous sodium dodecyl sulfate (SDS, Sigma, 9 mM) at a concentration slightly above the CMC (which is 8.5 mM) of SDS to give a viscous solution of concentration 0.5 mg of SWNT per milliliter of SDS.[19] The analytes were then added to the SWNT SDS solutions as outlined below.

A number of different SWNTs purification methods were attempted. All purified SWNT samples used to collect data reported in this paper were purified by refluxing in 3 M nitric acid at 120 °C for 13.5 h and then washed with water (18.2 MΩ). Purity was investigated by transmission electron microscopy (TEM).[4]

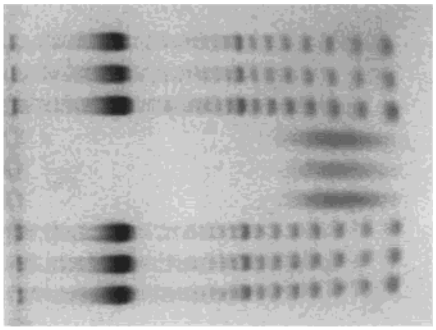


Figure 1. Gel electrophoresis analysis of a DNA sample after sonicating for about 2−3 h using 100 base pair ladder (where bands in the top and bottom three lanes are from right to left 100, 200, 300, 400, 500 bp, etc.). The middle three lanes contained the sonicated DNA sample to be used in the fragmented DNA−carbon nanotube experiments. The sonicated DNA sizes range from 200 to 400 base pairs.

All solvents (BDH laboratories) were analytical grade except for hexane, which was spectroscopic grade. Calf thymus DNA (ct-DNA), naphthalene, and anthracene were obtained from Sigma Aldrich Chemical Co. and used as received. A stock solution of ct-DNA (0.4 mg/mL) was prepared in SDS (9 mM).

All spectroscopy experiments except those noted as being in hexane were performed with aqueous SDS (9 mM) as the solvent. For the spectroscopy experiments, in the first instance a potential SWNT ligand was added to a SWNT suspension (∼0.1 mg mL^{-1}) in SDS to a concentration of ∼0.5 mg mL^{-1} for the small molecules and 0.1 mg mL^{-1} for the DNA, though in some cases further dilution (with SDS (9 mM)) was required to avoid excessive absorbance. Anthracene was introduced to the single-walled suspension either by adding the solid and vortexing the sample to facilitate its solubilization or by adding a concentrated methanolic solution to the SWNT suspension. Other polyaromatic hydrocarbons were introduced from methanolic solutions. Short DNAs of ∼200−400 base pairs (as determined by gel electrophoresis, Figure 1) were obtained by sonicating an aqueous SDS solutions of ct-DNA for about 2−3 h,[20] and aliquots were added to the SWNT suspension. All preparations were left overnight to equilibrate before spectroscopic measurements were performed. The DNA/SWNT samples gave the same results if the spectra were measured immediately upon mixing; however, leaving the samples for days resulted in some precipitation from solution and unsatisfactory spectroscopy. For the same reason the SWNT stock solution was always prepared fresh as SWNTs are not stable in SDS for more than 1 week because they aggregate together and fall out of solution. SDS (9 mM) was used as the baseline for all spectra; however, for the LD it was found that the sample without rotation could be used as the baseline because it overlaid on the rotating SDS spectrum. In fact, this coincidence of the baselines was a good check on system performance.

Spectroscopy. Absorbance: UV−visible absorbance spectra were recorded using a Cary 1E spectrophotometer.

Linear dichroism (LD): LD is the difference in anisotropic absorption of light polarized in planes parallel ($A_{||}$) and perpendicular ($A_{\perp}$) to the direction of orientation.[21]

$$\mathrm{LD} = A_{||} - A_{\perp} \tag{1}$$

An isotropic collection of molecules will have no LD, while a group of molecules whose transition moments are macroscopically oriented will absorb differing amounts of parallel and perpendicular plane-polarized light, producing a nonzero LD spectrum. For a perfectly

(10) Marconi, G.; Mayer, B.; Klein, C. T.; Kohler, G. *Chem. Phys. Lett.* **1996**, *260*, 589−594.
(11) Bommeli, F.; Degiorgi, L.; Wachter, P.; Bacsa, W. S.; deHeer, W. A.; Forro, L. *Solid State Commun.* **1996**, *99*, 513−517.
(12) Tsang, S. C.; Guo, Z. J.; Chen, Y. K.; Green, M. L. H.; Hill, H. A. O.; Hambley, T. W.; Sadler, P. J. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2198−2200.
(13) Buzaneva, E.; Karlash, A.; Yakovkin, K.; Shtogun, Y.; Putselyk, S.; Zherebetskiy, D.; Gorchinskiy, A.; Popova, G.; Prilutska, S.; Matyshevska, O.; Prilutskyy, Y.; Lytvyn, P.; Scharff, P.; Eklund, P. *Mater. Sci. Eng., C: Biomimetic Supramol. Syst.* **2002**, *19*, 41−45.
(14) Matyshevska, O. P.; Karlash, A. Y.; Shtogun, Y. V.; Benilov, A.; Kirgizov, Y.; Gorchinskyy, K. O.; Buzaneva, E. V.; Prylutskyy, Y. I.; Scharff, P. *Mater. Sci. Eng., C: Biomimetic Supramol. Syst.* **2001**, *15*, 249−252.
(15) Dovbeshko, G. I.; Repnytska, O. P.; Obraztsova, E. D.; Shtogun, Y. V. *Chem. Phys. Lett.* **2003**, *372*, 432−437.
(16) Nakashima, N.; Okuzono, S.; Murakami, H.; Nakai, T.; Yoshikawa, K. *Chem. Lett.* **2003**, *32*, 456−457.
(17) Guo, Z. J.; Sadler, P. J.; Tsang, S. C. *Adv. Mater.* **1998**, *10*, 701−703.
(18) Journet, C.; Maser, W. K.; Bernier, P.; Loiseau, A.; delaChapelle, M. L.; Lefrant, S.; Deniard, P.; Lee, R.; Fischer, J. E. *Nature* **1997**, *388*, 756−758.
(19) Doorn, S. K.; Fields, R. E.; Hu, H.; Hamon, M. A.; Haddon, R. C.; Selegue, J. P.; Majidi, V. *J. Am. Chem. Soc.* **2002**, *124*, 3169−3174.
(20) Elsner, H. I.; Lindblad, E. B. *DNA−J. Mol. Cell. Biol.* **1989**, *8*, 697−701.
(21) Rodger, A.; Nordén, B. *Circular Dichroism and Linear Dichroism*; Oxford University Press: New York, 1997.

oriented molecule the measured LD would equal $A_{\parallel}$ (>0) for a transition polarized exactly parallel to the orientation direction or $-A_{\perp}$ (<0) for a transition polarized exactly perpendicular to a uniaxial orientation direction. Thus, qualitative information about the orientation of molecules in space can be extracted from the sign of the LD. A more quantitative picture of molecular alignment can be obtained from the reduced LD (LD^r)[21]

$$LD^r = \frac{LD}{A} = \frac{3}{2}S(3\cos^2\alpha - 1) \qquad (2)$$

where A is the normal isotropic absorption and constitutes an average over all possible orientations, α is the angle between the transition moment of interest and the macroscopic orientation axis, and S is the orientation parameter (1 for perfect orientation and 0 for an isotropic solution).

Flow LD spectra were recorded using a Jasco J-715 circular dichroism spectropolarimeter with extended sample compartment adapted for LD measurements. The Couette cell used to flow orient the samples in most of the experiments has a CaF_2 rotating central cylinder and CaF_2 windows in an outer stationary sleeve separated by a gap of 50 μm, making a path length of 100 μm.[22] The photomultiplier tube was moved into the sample compartment next to the LD cell using a housing fabricated for the instrument by European Chirality Services[23] to reduce artifacts due to scattered light. Some experiments were performed with a small volume quartz LD cell whose outer quartz cylinder has an internal diameter of ~3 mm and rotates while the inner cylinder is a stationary ~2.5 mm diameter quartz rod.[24] The latter cell had the advantages of having focusing optics after the sample to collect scattered light and a demountable sample holding component which facilitates cleaning. The PMT was left in the standard position with this cell. The rotation speed used in the experiments was ~1000 rpm.

Gel Electrophoresis. A stock solution of Tris borate buffer (0.45 M) was prepared by dissolving Tris(hydroxymethyl)−aminomethane (27 g, 0.1 M, Amersham Pharmacia Biotech) and orthoboric acid (13.75 g, 0.1 M, BDH) in distilled water (500 mL). The pH of the buffer was adjusted to pH 7 with concentrated hydrochloric acid. On the day of the experiment a Tris borate buffer working solution (0.045 M) was prepared by diluting Tris borate buffer stock solution (50 mL, 0.45 M) to 500 mL with distilled water.

A 2% w/v agarose gel was prepared by heating agarose (0.67 g, Amersham Pharmacia Biotech) in Tris borate buffer working solution (33.5 mL, 0.045 M) to 95 °C with constant stirring until all agarose had dissolved. The solution was allowed to cool to 60 °C and then cast onto a gel tray (110 × 100 mm) with an 11-tooth comb used to produce sample wells already in position (depth of gel ≈ 4 mm). The gel was allowed to dry for 1 h at room temperature. The gel tray was then positioned in a Pharmacia GNA-100 submarine tank and just submerged under Tris borate buffer working solution. The comb was removed, and 6 μL of each sample was loaded into the individual wells.

A Pharmacia Electrophoresis power supply ECPS 3000/150 was used. The power supply was initially set to ~70 V and turned off when the markers had moved 3/4 of the way up the gel. The gel tray was removed from the tank and placed in a glass dish containing ethidium bromide (40 μL of 10 mg/mL, Amersham Pharmacia Biotech in 100 mL of Tris borate buffer working solution). The gel was stained for 15 min and then visualized under a UV lamp (at 254 nm) and photographed using a UVP White/UV transilluminator. Photographic images were obtained using Grab-it 2.0 software (Synoptics Ltd).

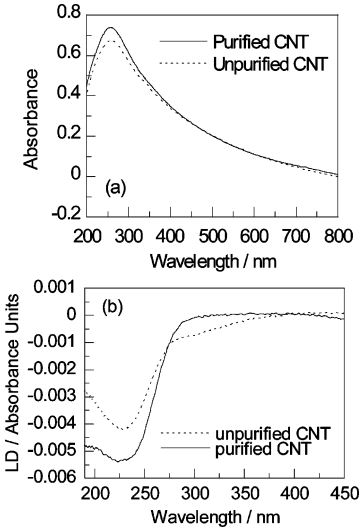


Figure 2. (a) Absorbance spectrum and (b) flow LD spectrum of unpurified (dashed line) and purified (solid line) carbon nanotubes (0.1 mg/mL) dissolved in 9 mM SDS. All spectra have had baselines of SDS (9 mM) without SWNTs subtracted.

Transmission Electron Microscopy (TEM). A carbon-coated copper mesh screen, 2−3 nm in diameter, was used to hold the samples for TEM.[25] A drop of the sample was spread over the carbon-coated copper grid in a thin film of material less than 100 nm thick. The grids were placed on filter paper to remove the excess material and dried at room temperature for 1 day, and then the image was captured. TEM images were obtained using a JEOL 2000FX transmission electron micrograph with a beam accelerating voltage of 200 kV.

3. Results and Discussion

The lack of information about the UV spectroscopy of CNTs in the literature probably correlates with the fact that the spectroscopy observed there is quite broad and in most spectrometers would be dominated by the scattering of light by the CNTs.[26] In our case we have either moved the photomultiplier tube very close to the LD cell or used a small volume LD cell with focusing lenses before and after the sample to ensure that most of the unabsorbed light is detected rather than scattered in our LD experiments.

Linear Dichroism (LD) of Nanotubes. The aim of this work was to determine if LD could be used to probe the structures of nanotube/ligand complexes. LD is measured on systems that are either intrinsically oriented or oriented during the experiment. Figure 2 shows LD spectra of unpurified and purified SWNTs, from which it is evident that the carbon nanotubes can be oriented by Couette flow. The signal is not due simply to light scattering or turbidity dichroism (i.e., loss of photons rather than absorbance), which would result in a positive LD signal as well as a positive absorbance with a (wavelength)$^{-n}$ dependence.[27]

(22) Rodger, A.; Rajendra, J.; Marrington, R.; Ardhammar, M.; Nordén, B.; Hirst, J. D.; Gilbert, A. T. B.; Dafforn, T. R.; Halsall, D. J.; Woolhead, C. A.; Robinson, C.; Pinheiro, T. J. T.; Kazlauskaite, J.; Seymour, M.; Perez, N.; Hannon, M. J. *Phys. Chem. Chem. Phy.* **2002**, *4*, 4051−4057.
(23) Rodger, A.; Rajendra, J.; Marrington, R.; Mortimer, R.; Andrews, T.; Hirst, J. B.; Gilbert, A. T. B.; Marrington, R.; Halsall, D.; Dafforn, T.; Ardhammar, M.; Nordén, B.; Woolhead, C. A.; Robinson, C.; Pinheiro, T.; Kazlauskaite J.; Seymour, M.; Perez, N.; Hannon, M. J. Flow oriented linear dichroism to probe protein orientation in membrane environments. In *Biophysical Chemistry: Membranes and Proteins;* Templer, R. H., Leatherbarrow, R., Eds.; Royal Society of Chemistry: Cambridge, 2002; pp 3−19.
(24) Marrington, R.; Dafforn, T. R.; Halsall, D. J.; Mortimer, R.; Andrews, T.; Rodger, A. Micro Volume Couette Flow Sample Orientation for Absorbance and Fluorescence Linear Dichroism. *Biophys. J.* **2004**.
(25) Sawyer, L. C.; Grubb, D. *Polymer microscopy*, 2nd ed.; Chapman and Hall: New York, 1995; pp 17−33.
(26) Kataura, H.; Kumazawa, Y.; Maniwa, Y.; Umezu, I.; Suzuki, S.; Ohtsuka, Y.; Achiba, Y. *Synth. Met.* **1999**, *103*, 2555−2558.
(27) Mikati, N.; Nordh, J.; Nordén, B. *J. Phys. Chem.* **1987**, *91*, 6048−6055.

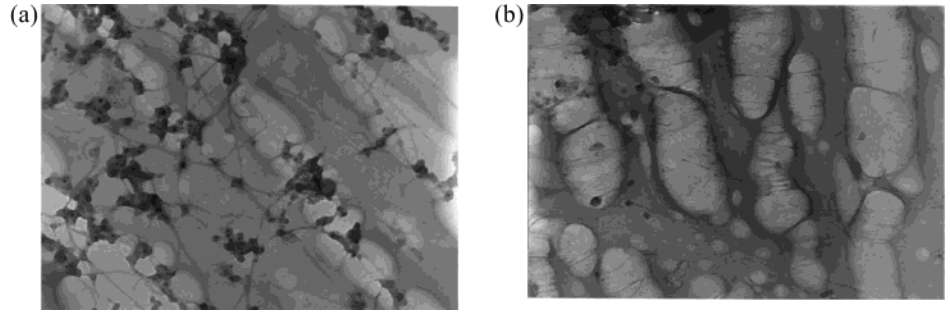


Figure 3. Transmission electron microscopy (TEM) images at ×25k magnification of (a) SWNTs and (b) SWNTs purified by refluxing in 3 M hydrochloric acid at 120 °C for 13.5 h.

In this case, we observe both a maximum in the absorbance spectrum and a negative maximum in the LD spectra at ∼225 nm. It is interesting to note that the purified sample has no LD signal above 275 nm, whereas the unpurified sample does. In addition, the signal for a given mass/volume concentration is larger for the purified sample below 275 nm. Thus, the purification process removes the impurities that absorb above 270 nm. Since the surfaces of single-walled nanotubes produced by catalytic arc discharge are normally coated with amorphous carbon or fullerene materials in which the metal catalyst remains embedded,[28] the differences between the unpurified and purified spectra suggest that the long wavelength signal is due to these materials and that the lower wavelength signal is intrinsic to the CNT. In support of this conclusion, TEM images (Figure 3) show that most of the surface material has indeed been removed by the purification process. The 225 nm region spectrum is presumably due to in-plane polarized $\pi-\pi^*$ transitions. That the LD is negative requires the polarization of this transition to be less than 54° from the average nanotube axis. Whether different geometries of SWNTs would have different LD signs at 225 nm or the expected long-axis polarization is at shorter wavelength are questions that cannot yet be answered. However, with improved nanotube production techniques the former question could be answered. Determination of the LD sign below 200 nm will require a further reduction in the light scattering contribution to the LD. This work is in progress.

Nanotubes and Small Aromatic Molecules. Given that the SWNT were oriented in the Couette flow, we were interested in seeing whether LD could be used to probe the interaction of small aromatic molecules with the nanotubes. A chromophore specifically bound to an oriented nanotube will also be oriented and is expected to show an LD signal under its own absorption bands. Our preliminary experiments showed signals for most of the molecules investigated; however, the signals were small, so we chose to focus on anthracene and naphthalene which have very distinctive spectroscopic signatures. These molecules are thus good tests for proof of principle and method development.

Anthracene. The LD spectrum of Figure 4c shows that anthracene binds to the SWNTs since anthracene has no intrinsic LD signal of its own in SDS solution. The 380 nm region of the anthracene spectrum is intrinsically polarized along the short axis of the anthracene molecule, though it has a long-axis component 'borrowed' from the 250 nm band. Polyethylene film LD of anthracene shows it to be dominated by the short-axis component.[21] That it has a negative signal tells us that the short

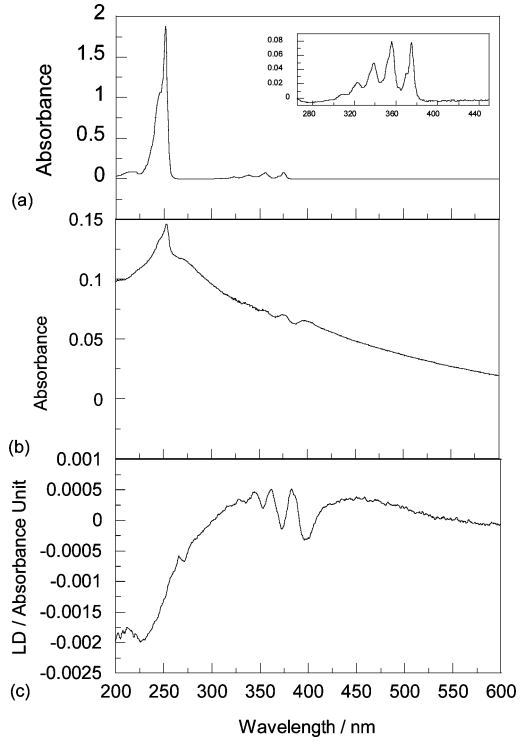


Figure 4. (a) Absorption spectrum of anthracene (0.0125 mg/mL, 1 mm path length) dissolved in hexane. The spectrum from 265 to 450 nm is enlarged in the insert. (b) Absorption (in 1 mm path length) and (c) LD spectrum of SWNT/anthracene complex. SWNT (0.025 mg/mL) in aqueous SDS (9 mM) by sonication was added to concentrated anthracene (0.25 mg) in methanol; baselines of, respectively, the absorbance and LD of SDS (9 mM) were subtracted in each case.

axis of anthracene is oriented at an angle larger than 54° with respect to the long axis of the SWNT. The shift to significantly longer wavelength than that of anthracene in hexane suggests a significant degree of $\pi-\pi$ interactions between the anthracene and the SWNT or between anthracene molecules. The 250 nm region of the anthracene spectrum (a long-axis-polarized transition) is difficult to probe in the presence of SWNT due in part to the broad CNT absorbance (see above) in this region. The LD has a small indication of a positive band in the region of 265 nm and perhaps a negative signal at ∼275 nm overlaid on top of the broad CNT LD (Figure 4c). This is significantly red shifted compared with the hexane wavelength of 250 nm and corresponds to a small broad shoulder in the CNT plus anthracene absorbance (Figure 4b). The 250 nm sharp band in the absorbance (Figure 4b) correlates with a free anthracene concentration of ∼0.002 mg/mL (total anthracene concentration is 0.0125 mg/mL). The most puzzling aspect of this spectrum

(28) Zhang, Y.; Shi, Z.; Gu, Z.; Iijima, S. *Carbon* **2000**, *38*, 2055−2059.

is that the 380 nm LD band has a significantly larger signal than the 265 nm band despite the absorbance intensity of this band being ~20 times smaller in the hexane absorbance spectrum (Figure 4a). However, absorbance of anthracene in the presence of the SWNTs shows the same effect, suggesting that some electronic interaction is taking place between the anthracene and SWNT.

The most obvious binding mode of the anthracene to the CNT is lying flat on the surface. If an anthracene molecule is indeed oriented flat on the surface of a nanotube oriented parallel to the long axis of the nanotube, then the 265 nm region transition has $\alpha = 0$ and the 380 nm band has $\alpha = 90°$ (from eq 2), giving, respectively, positive and negative signals (as observed), but the 380 nm region should be 1/40 the magnitude of the 260 nm region.

An alternative explanation is that the anthracene is oriented with the short axis perpendicular to the surface of the SWNT and there is no preferred orientation on the nanotube. The symmetry of the situation is then analogous to that of a ligand binding to a flow-deformed liposome and we may write[22,29]

$$\mathrm{LD}^{\mathrm{r}} = \frac{3S}{4}(1 - 3\cos^2\beta) \qquad (3)$$

where β is the angle the transition moment of interest makes with the normal to the nanotube surface. Transition moments perpendicular to the surface (380 nm region short-axis polarized here) thus have $\mathrm{LD}^{\mathrm{r}}/3S = -1/2$. If the average angle of the long-axis moment lies at an angle close to 54° from the CNT long axis, then we would predict LD signals of similar magnitude to those observed. However, the broadening and apparent loss of magnitude of the 250 nm absorbance signal is still not accounted for.

To investigate the role of packing effects on the observed LD spectra, three different concentrations 0.01, 0.02, and 0.05 mg/mL, corresponding to 10%, 20%, and 50% of full coverage, respectively (if all anthracene molecules are in solution—which is unlikely to be the case—and lie flat on the surface of the nanotube), of anthracene were chosen to analyze the effect of loading on the SWNTs (at concentration 0.1 mg/mL in SDS (9 mM)). It was not possible to measure accurate data at lower loadings, and the SWNT concentration could not be increased. The 250 nm anthracene absorbance is essentially lost, indicating little or no free anthracene in these preparations.

Naphthalene. The absorbance and LD spectra of naphthalene and SWNTs are shown in Figure 6 at reasonably high naphthalene concentrations. The LD spectrum of Figure 6c shows that naphthalene binds to the SWNTs since it has no intrinsic LD signal of its own in SDS solution. The naphthalene absorbance loses its structure upon interaction with the SWNT and shifts to longer wavelengths. The longer wavelength short-axis-polarized band LD, in contrast to the situation with anthracene, has an LD signal smaller than the long-axis polarized 220 nm region, consistent with its smaller absorbance magnitude. After subtracting the nanotube LD signal from the naphthalene/SWNTs signal (Figure 6d), there is a positive/negative couplet with a sharp negative LD signal at 218 nm and a broad longer wavelength positive signal. The LD baseline

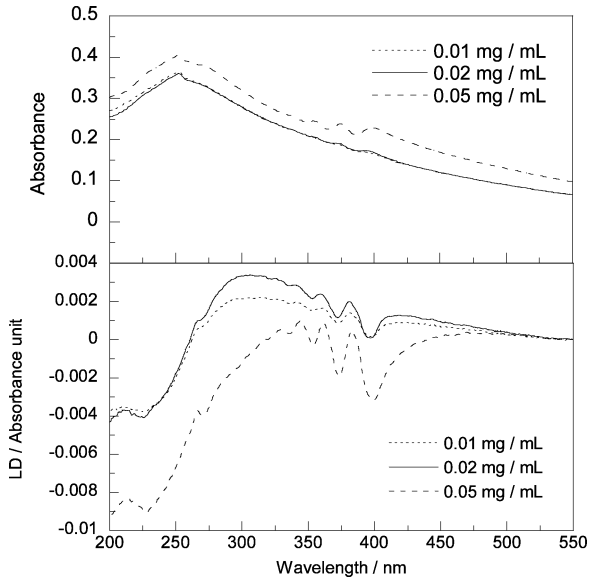


Figure 5. (a) Absorption and (b) linear dichroism spectra of different concentrations of anthracene with SWNTs (0.1 mg/mL)/SDS (9 mM). Anthracene was added to the SWNT solution as a concentrated methanolic solution, making final concentrations of 0.01, 0.02, and 0.05 mg/mL anthracene solutions. Baselines of, respectively, the absorbance and LD of SDS (9 mM) were subtracted in each case.

of Figure 6d is flat as we would normally expect after subtracting the intrinsic SWNT contribution. This is in contrast to the situation with anthracene, which argues for some kind of absorbance intensity transfer from the anthracene 250 nm band to the SWNT.

To probe the effect of surface coverage of the SWNT by naphthalene, data were collected for different concentrations of naphthalene (Figure 7). We assumed that the naphthalene absorbance coefficient is unchanged by its physisorption and used this extinction coefficient to calculate the amount of naphthalene in solution. The ratio of absorbance of naphthalene (Figure 6a) in a solution of known concentration to that seen in the difference spectrum of Figure 7b then gives the concentration of adsorbed naphthalene in a solution of SWNT/naphthalene complex.[7] The coverage of the following concentrations of naphthalene 0.2, 0.4, and 0.6 mg/mL were then calculated as 1.4%, 15%, and 26%, respectively, assuming the naphthalene lies flat on the surface of the SWNT (any absorbance suppression will mean these numbers are an underestimate). Though the signal-to-noise ratio is significantly worse at the lower loadings, from Figure 7d it is clear that the spectral shape depends significantly on the coverage with the 220 nm couplet not being evident at the lowest loading. Intriguingly, the long wavelength band is comparatively more intense at the lower loadings, as in the case of the anthracene spectrum (whose lower solubility precludes higher concentrations being studied). This leads one to speculate that the loss of the higher energy absorbance and LD intensity at lower loadings in both cases is due to some coupling between the nanotube and naphthalene or anthracene chromophore. This effect is stronger for anthracene, where a flat baseline could not be obtained by subtracting the absorbance of free SWNT.

At the lower loadings, both naphthalene bands have positive LD signals. A simple polyethylene stretched film LD experiment shows that both bands are dominantly long-axis polarized (data not shown). (Even though the electronic origin of the first band

(29) Ardhammar, M.; Lincoln, P.; Nordén, B. *J. Phys. Chem. B* **2001**, *105*, 11363−11368.

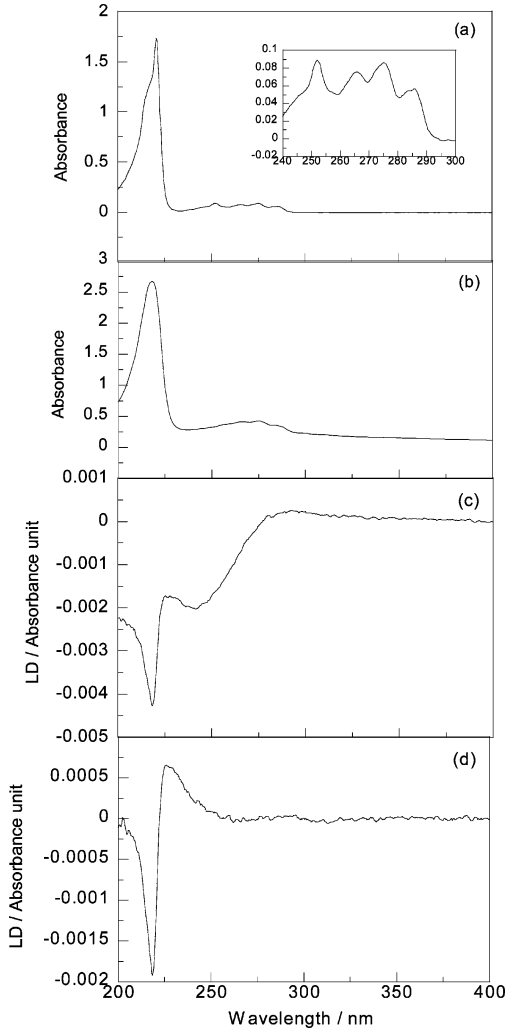


Figure 6. (a) Absorption spectrum of naphthalene (0.017 mg/mL) dissolved in hexane with 1 mm path length. The spectrum from 240 to 300 nm is enlarged in the insert. (b) Absorption and (c) LD spectra of SWNT/naphthalene complex. Naphthalene was added to the SWNT solution (0.1 mg/mL in 9 mM SDS) as a concentrated methanolic solution, making a final concentration of 0.5 mg/mL naphthalene solution. Baselines of, respectively, the absorbance and LD of SDS (9 mM) were subtracted in each case. (d) LD signal of naphthalene after subtracting the LD signal of SWNTs from SWNT/naphthalene signal

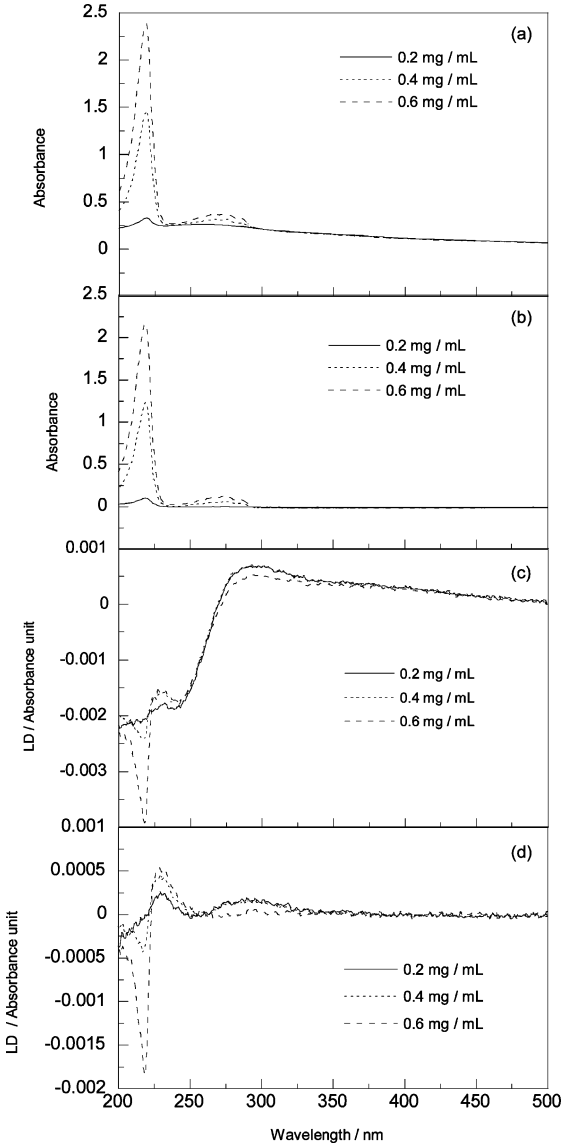


Figure 7. (a) Absorption spectra of different concentrations of naphthalene in SWNTs (0.1 mg/mL)/SDS (9 mM) solution. SWNTs/naphthalene complexes were prepared by adding different amounts (0.2, 0.4, and 0.6 mg) of methanol dissolved naphthalene to 1 mL of SWNTs (0.1 mg/mL) in SDS (9 mM) solution to make final concentrations of 0.2, 0.4, and 0.6 mg/mL naphthalene solutions. (b) Difference absorption spectra of part (a) spectra obtained by subtracting the absorbance spectrum of SWNT (0.1 mg/mL) from the absorbance spectra of 0.2, 0.4, and 0.6 mg/mL naphthalene/SWNTs complexes. (c) LD spectra of different concentrations of naphthalene in SWNTs (0.1 mg/mL)/SDS (9 mM) solution. SWNTs/naphthalene complexes were prepared by adding different amounts (0.2, 0.4, and 0.6 mg) of methanol-dissolved naphthalene to 1 mL of SWNTs (0.1 mg/mL) in SDS (9 mM) solution and making final naphthalene concentrations of 0.2, 0.4, and 0.6 mg/mL. (d) Difference LD spectrum of 0.2, 0.4, and 0.6 mg/mL naphthalene obtained by subtracting the LD spectrum of SWNT (0.1 mg/mL) from the LD spectra of 0.2, 0.4, and 0.6 mg/mL naphthalene/SWNTs complexes.

is short-axis polarized, the intensity is dominated by vibronic borrowing from the allowed shorter wavelength band). Thus, the naphthalene long axis is oriented at an angle smaller than 54° from the SWNT long axis than perpendicular to it. The data are consistent with the naphthalene lying flat on the surface but with no other orientation preference. The decrease in the positive signal of the longer wavelength band and the appearance of a sharp negative signal at 218 nm as the loading is increased is consistent with the naphthalene gradually adopting a mode with the long axis perpendicular to the SWNT long axis, i.e., poking out from the SWNT, and loss of coupling of these transitions to the electronic structure of the SWNT. The excitonic nature of this band is consistent with the degenerate coupling between transitions of π-stacked naphthalene units.

Nanotubes and DNA. Scanning tunneling microscopy (STM), transmission electron microscopy (TEM), and atomic force microscopy (AFM) are the main methods currently used for the visualization of DNA.[17] TEM images have been used to show that DNA binds to carbon nanotubes; still the nature of such interaction is not clear.[17] We were interested in investigating the ability of LD to probe the interaction of DNA and nanotubes in aqueous solution. Figure 8a shows that the sonicated DNA samples were not long enough to show significant orientation in our experiments. As shown in Figure 8d, it is clear that the DNA is binding to carbon nanotubes because the DNA-associated carbon nanotube gives a larger LD signal than the sum of the LD spectrum of DNA and the SWNT in isolation. Thus, the DNA is binding to the SWNTs and is

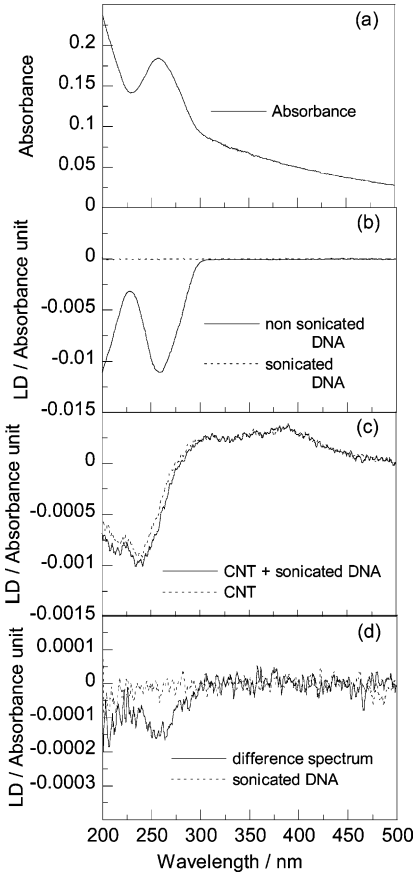


Figure 8. (a) Absorbance spectrum of sonicated DNA (0.033 mg/mL)/SWNTs (0.033 mg/mL) complex in SDS (9 mM) solution. The nonsonicated DNA absorbance overlays the sonicated absorbance at the same concentration. (b) LD spectrum of unsonicated (solid line) and sonicated (dashed lines) ct-DNA in aqueous SDS (9 mM). (c) LD spectra of SWNT (0.033 mg/mL)/sonicated DNA (0.033 mg/mL) (solid line) and LD spectrum of SWNT (0.033 mg/mL) (dashed line). (d) Difference LD spectrum of SWNT (0.033 mg/mL)/sonicated DNA (0.033 mg/mL) complex and SWNT (0.033 mg/mL) (solid line) spectrum compared with sonicated DNA (0.033 mg/mL, dashed line) spectrum. All spectra had SDS (9 mM) baselines subtracted.

being flow oriented when the SWNTs are oriented. The small size of the signal suggests very little is binding (not entirely surprising given the anionic SDS required to solubilize the CNTs) and/or the DNA is wrapped about the nanotube at an oblique angle, thus having the bases oriented fairly close to the magic angle of 54.7°.

Visual confirmation that the DNA was indeed binding to the SWNTs was obtained using TEM (Figure 9), which also suggests that DNA molecules tend to cover the surface of the nanotubes evenly. Platinated and iodated DNA oligomers and

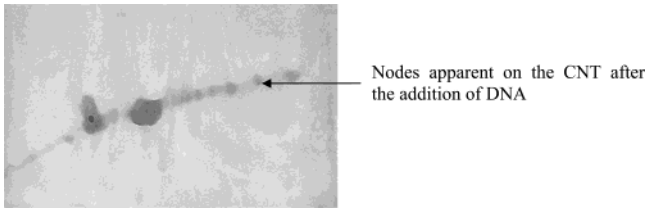


Figure 9. TEM images at ×40k magnification of DNA associated carbon nanotubes.

their immobilization on CNT have been visualized under high-resolution TEM;[12] however, we were unable to improve the resolution of our unsubstituted DNA. Nodes (indicated by an arrow) are apparent on the CNT, showing that DNA is binding to CNTs (no nodes are apparent for CNT images in the absence of DNA, Figure 3). In this image we could also see the amorphous carbon attached to the carbon nanotubes.

4. Conclusion

Three ligands, anthracene, naphthalene, and sonicated calf thymus DNA, were chosen to establish the potential of flow linear dichroism to probe ligands binding to single-walled carbon nanotubes. In each case clear evidence of binding was observed.

However, the data raise questions about the nature of the SWNT/ligand interactions especially in the case of the aromatic molecules where the spectroscopy of the ligands is significantly perturbed by their interaction with the nanotubes. The spectra for naphthalene are consistent with it adopting a mode where it lies flat on the surface at low loading, quite possible with no orientation preference or else aligned with the SWNT long axis, and at higher loadings it appears to begin adopting a mode with its long axis perpendicular to the surface. The anthracene experiments were more limited in concentration range, and our current proposal is that the anthracene is observed in a mode with its short axis perpendicular to the SWNT surface. The anthracene long-axis-polarized transition interacts significantly with the SWNT. In the case of naphthalene, when it lies flat on the surface of the SWNT the long-axis-polarized transitions are broadened and reduced in intensity. When naphthalene is oriented perpendicular to the surface, its long-axis-polarized transition is less perturbed.

A significant advantage of the LD, in addition to potentially providing orientation information, is that molecules free in solution are invisible so we only probe those bound to the CNTs.

Acknowledgment. Steve York's help with the electron microscopy is gratefully acknowledged.

JA048720J

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 47, Issue of November 19, pp. 48821–48829, 2004

Printed in U.S.A.

FtsZ Fiber Bundling Is Triggered by a Conformational Change in Bound GTP*

Received for publication, May 4, 2004, and in revised form, August 23, 2004
Published, JBC Papers in Press, August 23, 2004, DOI 10.1074/jbc.M404944200

Rachel Marrington‡§¶, Elaine Small§‖, Alison Rodger‡, Timothy R. Dafforn‡‡, and Stephen G. Addinall‖§§**

*From the ‡Department of Chemistry, University of Warwick, Coventry CV4 7AL, ‖School of Biological Sciences, Michael Smith Building, Oxford Road, Manchester M13 9PT, and **Department of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom*

Polymer formation by the essential FtsZ protein plays a crucial role in the cytokinesis of most prokaryotes. Lateral associations between these FtsZ polymers to form bundles or sheets are widely predicted to be extremely important for FtsZ function *in vivo*. We have carried out a study *in vitro* of FtsZ polymer formation and bundling using linear dichroism (LD) to assess structural properties of the polymers. We demonstrate proof-of-principle experiments to show that LD can be used as a technique to follow FtsZ polymerization, and we present the LD spectra of FtsZ polymers. Our subsequent examination of FtsZ polymer bundling induced by calcium reveals a substantial increase in the LD signal indicative of increased polymer length and rigidity. We also detect a specific conformational change in the guanine moiety associated with bundling, whereas the conformation and configuration of the FtsZ monomers within the polymer remain largely unchanged. We demonstrate that other divalent cations can induce this conformational change in FtsZ-bound GTP coincident with polymer bundling. Therefore, we present "flipping" of the guanine moiety in FtsZ-bound GTP as a mechanism that explains the link between reduced GTPase activity, increased polymer stability, and polymer bundling.

The FtsZ protein, a structural and functional homologue of eukaryotic tubulins, is a polymer-forming GTPase that plays a critical role in the cell division process in bacteria. Before the onset of cell division in the γ-proteobacterium *Escherichia coli*, FtsZ polymers form a dynamic ring (the Z-ring) around the cell center on the inner surface of the cytoplasmic membrane (1–3). The Z-ring recruits at least ten accessory proteins which are essential for cell division to proceed (4–7), and then reduces in diameter until cytokinesis is complete. FtsZ is extensively conserved among prokaryotes with only one phylum of archaebacteria (Crenarchaeota) and three free-living species of bacteria (*Ureaplasma urealyticum*, *Prosthecobacter dejongii*, and *Pirellula sp.*) proven to lack *ftsZ* genes thus far (8, 9). In addition, FtsZ is required for chloroplast division and may be essential for division of some mitochondria (10–12). Elucidating the mechanism by which the Z-ring drives constriction is, therefore, critical to our understanding of cell division in prokaryotes and cytokinesis in eukaryotic organelles.

Monomers of FtsZ protein associate *in vitro* in a GTP-dependent fashion into linear, unbranched, polymeric fibers whose dynamics result in the hydrolysis of GTP (13, 14). There is considerable evidence to suggest that the interface between two FtsZ monomers in the polymeric form is crucial for this GTPase activity, with the "top" and "bottom" of consecutive monomers combining to form the GTP-binding pocket (15–19) (Fig. 1). Although GTP binding favors polymer formation, the hydrolysis of GTP to GDP promotes shortening and bending of the FtsZ polymers (20). The simplest polymer form of FtsZ, a one-monomer wide protofilament, has been observed *in vitro* (21); however, FtsZ fibers are more commonly seen as pairs of parallel protofilaments called thick filaments *e.g.* Refs. 18, 21, and 22. More complex FtsZ oligomeric structures, which include sheets (or ribbons), tubes, and bundles, have also been observed *in vitro* (18, 22–29). However, in each case where detailed structure has been obtained, these consist of thick filaments in different arrangements (parallel, anti-parallel, twisted, straight, etc.) (18, 22, 24). Critically, the structure of the Z-ring *in vivo* remains uncharacterized; however, it seems likely from these *in vitro* studies that the Z-ring consists of thick filaments arranged in either sheets/ribbons or bundles.

FtsZ polymerization is dynamic both *in vivo* and *in vitro*. *In vivo* the Z-ring polymerizes quickly (30, 31) and undergoes constant turnover throughout division (31). *In vitro* measurements have determined that FtsZ polymer dynamics depend on multiple factors. These include pH, concentrations of magnesium, potassium, calcium, and competing nucleotide tri-phosphates, macromolecular crowding, and the presence of FtsZ-interacting proteins (23, 26, 27, 29, 32–37). Crucially, differences in the type of polymers formed under these various conditions correspond to differences in their longevity and dynamics. An example of this is the bundling of polymers caused by the presence of millimolar concentrations of calcium (27, 33). This bundling may be similar to that observed when FtsZ interacts with accessory proteins such as ZipA and ZapA *in vitro* (23, 28) and (at least in the case of calcium) results in a reduction of GTPase activity and slower turnover of monomers (27, 34). Subtle changes to the extent of bundling are also observed when FtsZ is polymerized in buffers of varying pH. Consistent with the effects of calcium, decreasing pH results in increased bundling, lower GTPase activity, and more stable polymers (26, 33). Therefore, controlling the extent to which lateral interactions occur between FtsZ fibers is likely to be important for Z-ring function *in vivo*.

* The work was funded in part by Biotechnology and Biological Sciences Research Council, Swindon, United Kingdom Grant REI120503 and Engineering and Physical Sciences Research Council Grant GR/M91105. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Both authors contributed equally to this work.
¶ Funded by Syngenta.
‡‡ A Medical Research Council Career Development Fellow. To whom correspondence should be addressed. Tel.: 44-121-414-5881; E-mail: T.R.Dafforn@bham.ac.uk.
§§ A Wellcome Career Development Fellow.

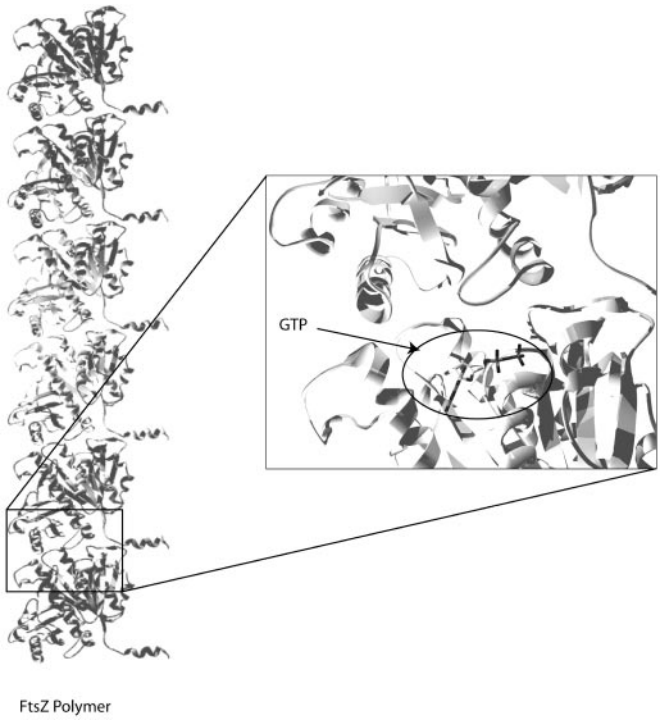


FIG. 1. *Left*, proposed structure of the FtsZ polymer fiber (18). The fiber axis runs vertically up the page. *Right,* an expanded view showing GDP (in *stick* representation) bound to FtsZ. The long axis of the guanine moiety extends out of the plane of the page, perpendicular to the fiber axis. The short axis is inclined to the fiber axis. The figure was produced using SwissPDB (59).

The alteration in lateral association caused by these extraneous factors is a process that is little understood at a structural level. It is likely that such associations are caused by alterations of the surface properties of the FtsZ polymer that encourage fiber-fiber interactions. It has been proposed that such changes occur by charge shielding of negatively charged FtsZ polymers (28). However, other mechanisms are possible. A conformational change within the FtsZ monomers making up the polymer or a change in the configuration of the monomer within polymers could each lead to exposure of inter-fiber binding sites. Alternatively, a combination of mechanisms may be responsible for the bundling of FtsZ fibers.

Techniques for examining solution phase conformational changes in proteins are fairly well established, the most commonly used being circular dichroism (CD) and Fourier transform infrared spectroscopy. Both provide spectroscopic information that can be deconvolved to provide measures of secondary structure content and, to some extent, side chain arrangement. Unfortunately, information on the orientation of monomeric protein units within protein polymers is less easy to obtain. In general, determining this sort of information has relied on either indirect biophysical measurement (*e.g.* using resonance energy transfer to triangulate monomer positions relative to one another (38)) or direct visualization of the polymer using electron microscopy (*e.g.* Ref. 18).

We have developed the technique linear dichroism (LD)[1] (which is technically related to CD) that provides information on the orientation of secondary structural elements and aromatic side chains within protein fibers. LD is the measure of the difference in absorbance by a sample of light polarized in orthogonal directions. To obtain an LD signal the fibers are usually aligned perpendicular to the incident light beam. We have achieved this by the use of a Couette flow cell which induces alignment as a result of shear flow in a liquid. Our own advances in cuvette design have only recently permitted the use of LD to examine relatively low volume (200 μl) biological samples (39, 40). Our initial studies of some common protein fibers, including actin, amyloid, and collagen, have allowed us for the first time to assign the signals observed to known secondary structure types (41). The study presented here represents a further advancement of the technique, with reduction of sample volume to 25 μl (42).

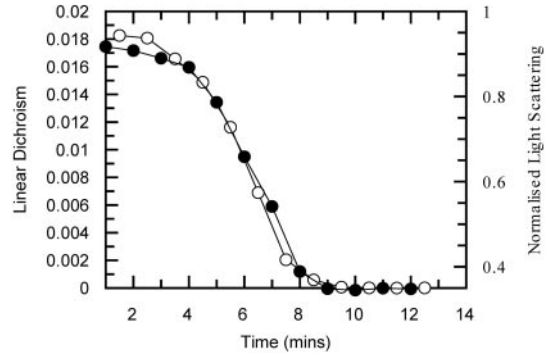


FIG. 2. **Polymerization of FtsZ at 11 μM with 0.2 mM GTP added at 0 s monitored by the change in LD signal at 210 nm (*open circles*) or light-scattering (*filled circles*).**

In this work we have used a combination of linear and circular dichroism to examine FtsZ polymerization and bundling. We demonstrate that LD provides a measure of FtsZ polymerization kinetics that is consistent with accepted results from light-scattering measurements. We show that bundling of FtsZ protofilaments induced by the presence of millimolar amounts of calcium is detectable as a substantial increase in LD signal. Importantly, we also detect a significant change in the linear dichroism signal in the guanine region of the spectrum under conditions where polymer bundles are formed. We also show that elevated levels of Mg^{2+} can mimic the bundling properties of Ca^{2+}. We have, therefore, detected a change in the conformation of the FtsZ-bound guanine moiety that correlates with (and probably accounts for) the decrease in GTPase activity associated with FtsZ polymer bundling. Hence, we provide the first putative mechanism for the link between increased polymer bundling and reductions in polymer dynamics and GTPase activity.

MATERIALS AND METHODS

FtsZ Purification—FtsZ was purified as described previously (36) using a modification of protocols described by Mukherjee and Lutkenhaus (43) and Lu and Erickson (44).

Spectroscopic Measurements—Light-scattering measurements were performed as described previously (36) with the exception of the direct side-by-side comparisons with LD for Fig. 2. These were performed at room temperature (23 °C) using a PerkinElmer Life Sciences LS50B spectrofluorimeter with excitation and emission wavelengths of 400 nm and a slit width of 2.5 nm with a 0.3-cm path length rectangular cuvette.

CD measurements were performed with a Jasco (Tokyo, Japan) J-715 spectropolarimeter at room temperature using a 0.1-mm path length demountable quartz cuvette. Data were collected every 0.5 nm from 260 to 185 nm, with a scanning speed of 100 nm min^{-1} and a response time of 1 s with data averaged over 8 scans.

LD measurements were performed at room temperature using a Jasco J-715 spectropolarimeter adapted for LD spectroscopy (42). Samples were aligned in the light beam using a custom made Couette cell. The cell was constructed by Crystal Precision Optics, Rugby, UK and consists of a cylindrical cross-section sleeve with a removable quartz capillary (sealed at one end with Araldite Rapid®) held centrally with respect to its circular face by a rubber O-ring. A quartz rod was suspended from the demountable lid into the capillary, creating an annular gap between rod and inner capillary wall of ~0.25 mm. The rod and the

[1] The abbreviations used are: LD, linear dichroism; MES, 4-morpholineethanesulfonic acid.

center of the capillary were aligned so that the capillary was able to rotate freely. The sample (25–60 μl) was placed into the capillary before loading into the cylindrical sleeve. The lid and rod units are then lowered into position. Upon rotation of the capillary (the rod remains stationary) a shear force is induced across the sample. The cylindrical sleeve has two windows for the light to pass through to make contact with the sample. This configuration allows the use of smaller sample volumes (25–60 μl) than has previously been possible (previous cells have required 200 μl (39, 40) or 2000 μl (45) of sample). The voltage applied to the motor that rotates the outer quartz cylinder is controlled electronically to allow the sample solution to be maintained with the highest possible degree of alignment without inducing turbulent flow and Taylor vortices. Data were collected using an interval scan measurement program available within the Jasco software. This enabled single full wavelength scans from 350 to 190 nm to be collected every minute at a scanning speed of 200 nm min^{-1}, data pitch 0.5 nm, and with a response of 0.5 s, thus, monitoring of the kinetics of polymerization across the whole wavelength spectrum. The spectra are of suitable intensity such that only one scan was necessary. Base lines of data collected after de-polymerization of FtsZ were subtracted from all spectra. The wavelength ranges of the LD spectra were in some cases restricted by the high absorbance of the samples at low wavelength. Data reported in this paper are truncated at a point where the Beer-Lambert law is valid. LD spectra were corrected for light-scattering as per Nordh *et al.* (46).

FtsZ polymerization assays were performed essentially as described previously (36). Briefly, FtsZ at the concentration specified in individual experiments was incubated in a standard polymerization buffer (50 mM MES, pH 6.5, 10 mM $MgCl_2$, 50 mM KCl), and GTP (the disodium salt adjusted to neutral pH) was added to a final concentration of 0.05–0.2 mM (see individual experiments) to initiate polymerization. Where specified, $MgCl_2$ was omitted from the standard buffer, or its concentration was altered. In some experiments $CaCl_2$ or $CuCl_2$ were added to the basic polymerization buffer to give the final concentrations described for individual experiments in the text and figure legends.

For all spectroscopic measurements samples were placed in the cuvette/capillary cell immediately upon preparation. Of the three techniques used, the longest dead time (about 40 s) was that taken to load and assemble the capillary LD unit and start the analysis.

GTPase Assay—GTPase activity of newly purified FtsZ protein was determined using γ-^{32}P-labeled GTP (Amersham Biosciences) and thin-layer chromatography as previously described (36). For titration of the effects of various ion concentrations on FtsZ GTPase activity, we measured production of NADPH from an enzyme-linked assay that has been used previously with FtsZ (47). This assay gave comparable values for FtsZ GTPase activity to the radioactive method but was more appropriate for multiple repeat experiments.

RESULTS

Measurement of FtsZ Polymerization Kinetics Using LD—For a linear dichroism signal to be measured, the molecule studied has to be aligned to some degree. In the case of the flow alignment Couette system used in these studies, the molecule is aligned by virtue of it having a shape with a high aspect ratio. We predicted that LD would be an ideal technique for studying FtsZ, which polymerizes from monomers (or short oligomers) into long, unbranched, linear polymers. Monomeric units of FtsZ (which roughly approximate to spheres and have an aspect ratio close to 1) will not align and, hence, will not produce an LD signal. Polymeric forms of FtsZ were expected to align and, thus, have a flow LD signal.

To test this prediction FtsZ polymerization reactions were followed by both LD and light-scattering. Samples to be observed by LD were placed in the LD cell, and their LD spectrum was measured in the UV, where the transitions of the chromophores of the protein backbone can be observed. GTP was then added to induce fiber formation, and the change in the LD signal was measured at a wavelength that is sensitive to the presence of aligned α-helical structure (FtsZ has a high α-helix content that should become apparent in the LD if fibers are formed and align). Control samples, also at room temperature, were studied by light-scattering as described under "Materials and Methods."

The measurement of the LD signal of FtsZ polymerization (Fig. 2) with time showed a very similar-shaped trace to that observed using right-angled light-scattering. It can be seen (Fig. 2) that there is an immediate large LD signal induced by the fast GTP-dependent association of FtsZ monomers to form polymers. There is then a plateau in the trace as the GTP in solution is turned over followed by a sigmoidal decrease as the GTP is exhausted and the polymers dissociate. The consistency between light-scattering and LD validates the use of LD to monitor the polymerization of FtsZ since the observed reaction kinetic is very similar to that observed previously by light-scattering (and other) techniques, *e.g.* Refs. 26, 33, 34, and 36.

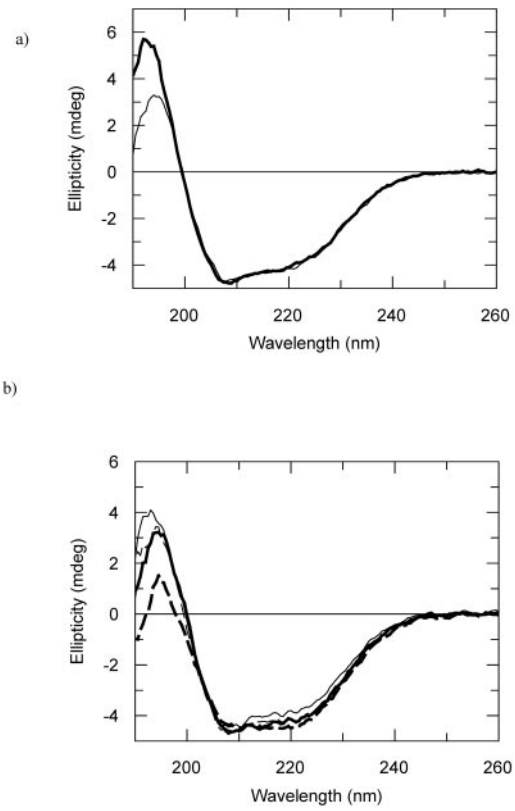


FIG. 3. **CD spectra of FtsZ at 11 μM in 50 mM MES buffer, pH 6.5, 10 mM $MgCl_2$, and 50 mM KCl at pH 6.5 at 23 °C alone (*thick line*) and in the presence of 0.2 mM GTP (*thin line*) (*a*) and in the presence of 0.2 mM GTP (*thick line*) and 1 mM (*thin line*), 5 mM (*dashed line*), and 10 mM (*thick dashed line*) Ca^{2+} (*b*).** Experiments were carried out in a 0.1-mm path length demountable cuvette.

CD Spectra of FtsZ and FtsZ with GTP—To understand the conformational changes that occur within FtsZ when the protein polymerizes, CD spectra were recorded of FtsZ in the presence and absence of GTP (Fig. 3*a*). These data show no discernable alteration in the signal collected above 200 nm as reported previously (48), with the only alteration between the two preparations being a reduction in signal intensity below 200 nm for the polymerized sample. Results from the deconvolution of the spectra of FtsZ using CDsstr (Table I) showed no significant differences from the secondary structure composition calculated from the x-ray crystal structure of FtsZ (Protein Data Bank code 1FSZ). The decrease in signal intensity of the CD spectra of FtsZ plus GTP compared with FtsZ alone can be attributed to light-scattering from the polymers that have formed upon addition of GTP.

CD Spectra of FtsZ and GTP with and without Ca^{2+}—Having established that only minor changes in FtsZ conformation occur upon GTP-dependent polymerization, we investigated whether bundling of FtsZ polymers was associated with conformational change in FtsZ. The presence of millimolar quantities of Ca^{2+} in standard FtsZ polymerization reactions has

TABLE I
Calculated secondary structure composition of FtsZ from deconvolution of CD data by CDsstr (60) compared to the secondary structure content calculated from the x-ray structure of 1FSZ using SwissPDB (59).

	Number of residues	α-Helix	β-Sheet	Other
		%	%	%
PDB file (1FSZ)	352	38	27	35
FtsZ	383	38	28	36
FtsZ + GTP	383	36	29	35
FtsZ + GTP + 5 mM Ca^{2+}	383	37	28	36

previously been shown to induce polymer bundling (27, 33). We, therefore, used CD to compare the backbone conformation of polymerized FtsZ in the absence of Ca^{2+} and in the presence of increasing concentrations of Ca^{2+} (1, 5, and 10 mM). It should be noted that any changes induced into the GTP CD upon protein binding will contribute to these spectra. The spectra from these experiments (Fig. 3*b*) showed very small alterations in the region between 200 and 250 nm, with a larger change occurring at wavelengths below 200 nm when the Ca^{2+} concentration was increased to 10 mM. Deconvolution of these spectra with CDsstr showed only small alterations in secondary structure upon the addition of increasing amounts of Ca^{2+} (Table I). The variation in CD signal intensity below 200 nm is due to light-scattering caused by the polymerization and subsequent bundling of FtsZ upon addition of Ca^{2+}.

LD Spectra of FtsZ and FtsZ with GTP—Having been unable to demonstrate significant changes in FtsZ conformation associated with either polymerization or polymer bundling by CD, we examined LD spectra of FtsZ under various conditions. We first obtained LD spectra of FtsZ in monomeric and polymeric form. The spectrum of the unpolymerized material is zero, as predicted, because molecular alignment is required to produce a signal (data not shown). Spectra from polymeric FtsZ (Fig. 4*a*) were obtained during the plateau phase of the reaction and have been truncated at a point where the Beer-Lambert law is followed. The ~220-nm n-π^* transition signal of the α-helix in the protein is apparent as a shoulder under the large lower wavelength signal due to the π-π^* transition (41). The sign of this transition is positive so its transition polarization is more parallel than perpendicular to the fiber axis. Because this transition lies perpendicular to the α-helix axis, this means that on average the helices within the protein lie more perpendicular than parallel to the fiber axis (39, 40).

At longer wavelengths (between 250 and 300 nm) LD can provide information on transitions of aromatic constituents within the FtsZ fiber. These features contain contributions from phenylalanine, tyrosine, and the guanine base of GTP (*E. coli* FtsZ has no tryptophan residues). Similar features have been observed for ATP-actin (41, 49) and can provide useful information on the inclination of these constituents to the fiber axis. In the case of the FtsZ spectrum in standard polymerization buffer (*i.e.* in the presence of 10 mM $MgCl_2$) the broad negative band across the aromatic region indicates that the planes of the chromophores lie more perpendicular than parallel to the fiber axis.

The Effect of Ca^{2+} on the Kinetics of FtsZ Assembly Measured by LD—We examined the LD spectra of FtsZ in the absence and presence of Ca^{2+}. First we used LD to compare the polymerization of FtsZ in the presence of increasing concentrations of Ca^{2+}. These data show two results of increased Ca^{2+} concentration; first, the length of the plateau phase of the polymerization is increased (Fig. 4*b*), and second, the amplitude of the peaks in the backbone region of the spectra were significantly increased (Fig. 4, *b* and *c*). The increase in plateau duration has been observed previously by light-scattering and corresponds to a decrease in GTPase activity (Refs. 27, 33, and 50 and Table II). The increase in the amplitude of the backbone region of the LD spectra (Fig. 4, *b* and *c*) is consistent with polymer bundling improving the degree of alignment of the polymer in the Couette cell because of an increase in length and/or rigidity of the polymers.

The Effect of Ca^{2+} on the Conformation of GTP Bound to FtsZ—We observed that the LD spectrum of the near UV region of polymeric FtsZ showed complex changes upon the addition of Ca^{2+} (Fig. 4*d*). Upon increasing the concentration of Ca^{2+} in the reaction mixture, the broad negative signal of the polymerized FtsZ changed to one with a negative signal at 280 nm and a larger positive signal at 250 nm. The sign change at 250 nm results from a steady increase in the intensity of a band at 250 nm, which moves from negative to positive. The chromophore that is responsible for this peak is the guanine of the bound GTP, which has absorbance maxima at ~250 and ~280 nm corresponding to an approximately short axis-polarized transition on the guanine at 278 nm and an approximately long axis-polarized transition at 244 nm (51, 52) (we note here that a recent study has conclusively demonstrated that more than 80% of the nucleotide bound to dynamic, polymeric FtsZ is GTP (13); therefore, we are confident that the overwhelming majority of the nucleotide that we are detecting by LD is GTP). The sign change that the guanine peak undergoes indicates that there is a significant structural rearrangement of the base, essentially moving it from an orientation where both the short axis and the long axis of the guanine are significantly more than 54.7° (the negative to positive change-over point for LD signals) from the fiber axis, to an orientation where the long axis is much closer to being parallel to the fiber axis, whereas the short axis is in a similar orientation to that in the calcium-free sample. Accounting for the fact that the backbone region of the LD (Fig. 4*c*) also shows an overall increase in fiber orientation with the addition of calcium and noting that the maximum of a positive LD signal can be twice that of a negative one (53), the long axis of the guanines can be seen to flip from an angle of $\sim(90 - \theta)°$ to the fiber axis to an angle of $\sim\theta°$ from the fiber axis, where $\theta \ll 35°$.

Flipping of the Guanine Moiety Can Be Induced by Both Ca^{2+} and Mg^{2+}—Because these experiments, along with all previous reports of FtsZ polymer bundling by Ca^{2+}, have related to conditions where both Mg^{2+} and Ca^{2+} were present, we sought to clarify the role of the different divalent cations in the guanine-flipping that we had observed upon FtsZ polymer bundling. Our standard polymerization conditions include 10 mM Mg^{2+} and no Ca^{2+} (see under "Materials and Methods"), and it has been demonstrated previously that polymerization of FtsZ does not require the presence of Ca^{2+} (33). Although Mg^{2+} is required for the FtsZ GTPase activity, it is not required for polymerization *per se*, since short, thin polymers are formed when GTP is added to FtsZ in the absence of Mg^{2+} (Ref. 26 and Fig. 5*a*). Therefore we examined GTP-dependent FtsZ polymerization in the presence of Ca^{2+} only. To rule out complications arising from the possible presence of residual amounts of Mg^{2+} ions in our purified FtsZ preparation, we added 0.1 mM EDTA to reaction mixtures before the addition of $CaCl_2$. Experiments were then examined by light-scattering, negative stain electron microscopy, and LD.

We find that FtsZ polymerization occurs in the presence of 10 mM Ca^{2+} but that the kinetics differ significantly from reactions containing 10 mM Mg^{2+}. In the presence of Ca^{2+} only, FtsZ polymerization occurs very slowly, such that the time taken to reach the plateau phase of the reaction (judged by light-scattering) is increased 100-fold compared with the same

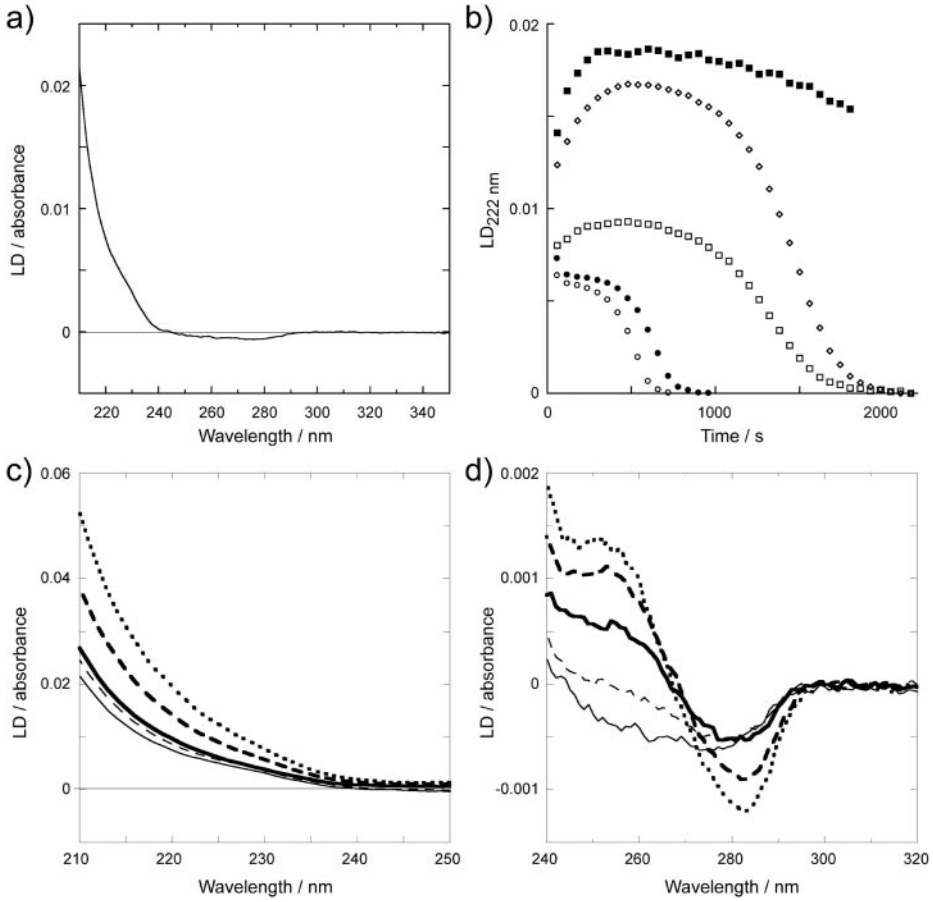


FIG. 4. *a*, the LD spectrum of FtsZ at 11 μM in 50 mM MES buffer, pH 6.5, 10 mM $MgCl_2$, and 50 mM KCl at 23 °C 1 min after the addition of 0.2 mM GTP. *b*, the change in LD at 222 nm of FtsZ as a function of time. Measurement was begun after the addition of 0.2 mM GTP to 11 μM FtsZ at 23 °C in 50 mM MES buffer, pH 6.5, 10 mM $MgCl_2$, and 50 mM KCl in the presence of 0 (*open circles*), 1 (*filled circles*), 3 (*open squares*), 5 (*open diamonds*), and 10 mM Ca^{2+} (*filled squares*). *c*, the effect of Ca^{2+} on the backbone region of LD spectra. *d*, the effect of Ca^{2+} on the aromatic region of LD spectra. Both *c* and *d* refer to FtsZ at 11 μM in 50 mM MES buffer, pH 6.5, 10 mM $MgCl_2$, and 50 mM KCl at 23 °C alone (*thin line*) and in the presence of 1 (*thin-dashed line*), 3 (*thick line*), 5 (*thick-dashed line*), and 10 mM Ca^{2+} (*dotted line*).

TABLE II
GTPase activity of FtsZ under various buffer conditions

The buffer contained 50 mM MES, pH 6.5, 50 mM KCl plus the indicated concentrations of $CaCl_2$ and $MgCl_2$.

$CaCl_2$	$MgCl_2$	GTPase activity
mM	*mM*	*mol/mol/min*
0	0	0.01
0	10	1.8
10	0	0.00
10	10	0.35
0	40	0.21

reaction in the presence of 10 mM Mg^{2+} (data not shown). Also, we were unable to detect hydrolysis of GTP by FtsZ polymers under these conditions (Table II). Despite this, FtsZ polymerization in the presence of Ca^{2+} and the absence of Mg^{2+} resulted in increased light-scattering (data not shown) and increased bundling of polymers (Fig. 5*d*) compared with the no calcium control (Fig. 5*a*). Crucially, we also observed changes in the LD spectrum indicative of polymer bundling, namely increased amplitude of the peaks in the backbone region and an inversion of signal at 250 nm (Fig. 6*a*), confirming that this is an ordered association and not aggregation. Therefore, 10 mM Ca^{2+} induces FtsZ polymer bundling with associated flipping of the guanine moiety despite being unable to support the catalytic GTPase activity normally associated with FtsZ polymerization.

Because the GTPase activity of FtsZ varies with the concentration of Mg^{2+} in the polymerization buffer (26), we decided to test whether increasing the Mg^{2+} concentration significantly above the optimum for GTPase activity could induce bundling and guanine-flipping as does 10 mM Ca^{2+}. At Mg^{2+} concentrations above 10 mM we observed increased light-scattering together with a decrease in polymer dynamics as assayed by light-scattering (Fig. 6*b*), and at 40 mM Mg^{2+} these properties were similar to those observed in our previous bundling experiments containing 10 mM Mg^{2+} and 10 mM Ca^{2+} (Fig. 6*b*). We also found that the GTPase activity of FtsZ in buffer containing 40 mM Mg^{2+} was reduced (Table II). We examined polymers formed under these conditions by electron microscopy and found increased amounts of bundling (Fig. 5*e*) compared with the 10 mM Mg^{2+} control (Fig. 5*b*). Finally we found that the LD spectra of FtsZ polymers formed in the presence of 40 mM Mg^{2+} were similar to those formed in the presence of 10 mM Mg^{2+} and 10 mM Ca^{2+}, exhibiting increased amplitude of the peaks in the backbone region and an inversion of signal at 250 nm diagnostic of the "flipped" configuration of the guanine moiety (Fig. 6*a*). In summary the addition of 30 mM extra Mg^{2+} to our standard FtsZ polymerization reaction reduced the GTPase activity ~10-fold, reduced polymer dynamics, induced polymer bundling, and induced a configuration change in the guanine of GTP from more perpendicular to the fiber axis to more parallel (Table II, Figs. 5 and 6). These are very similar consequences to the addition of 10 mM Ca^{2+} (Table II, Fig. 4).

DISCUSSION

We have observed that the LD spectra of FtsZ are influenced by increases in Mg^{2+} and Ca^{2+} concentration, leading to a significant increase in signal at 210 nm (Fig. 4). The LD magnitude increase suggests an increase in either the rigidity or length of the polymer (hence, increased alignment). Alternatively, the α-helices of the monomers could orient more perpendicular to the fiber axis or increase in number, but if this was the case (that is, Ca^{2+} induced a large change in the fold of the FtsZ monomer) it would be expected that the CD spectra of the backbone region would be altered significantly, which is not the case (Fig. 3). Thus, we observe a stiffening/lengthening of the fibers in the presence of Ca^{2+}. This correlates well with electron microscopy images of FtsZ polymers formed in the

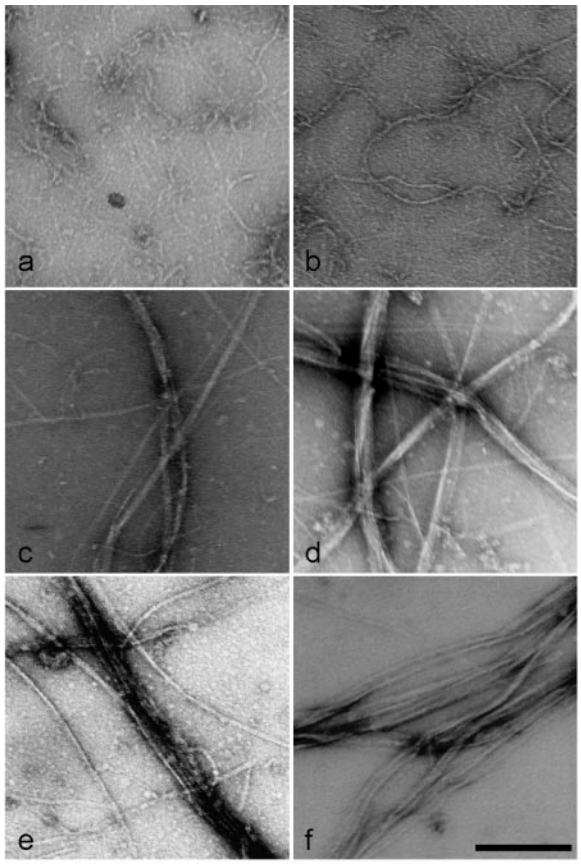


FIG. 5. **Morphology of FtsZ polymers.** FtsZ at 8.3 μM in 50 mM MES buffer, pH 6.5, and 50 mM KCl at 30 °C after the addition of 0.1 mM GTP in the presence of no divalent cations (*a*), 10 mM $MgCl_2$ (*b*), 10 mM $MgCl_2$ + 10 mM $CaCl_2$ (*c*), 10 mM $CaCl_2$ (*d*), 40 mM $MgCl_2$ (*e*), and 10 mM $MgCl_2$ + 0.1 mM $CuCl_2$ (*f*).

presence of 10 mM Mg^{2+} and 10 mM Ca^{2+} (Refs. 27 and 33 and Fig. 5*c*) or 40 mM Mg^{2+} (Fig. 5*e*), which, as well as forming frequent bundles, are demonstrably straighter in appearance than those formed in standard reactions (*e.g.* Refs. 33 and 36 and Fig. 5*b*).

Our LD data also show a significant change in one part of the near UV region in response to the presence of Ca^{2+} or elevated concentrations of Mg^{2+}. The inversion of a peak at 250 nm indicates a large movement of the guanine chromophore within the fiber, altering its orientation from having its long axis approximately perpendicular to approximately parallel to the fiber axis (strictly from $\gg$55° to $\ll$55°). This result elegantly demonstrates one of the major advantages of LD over other optical techniques; it can probe the orientations of key units within the fiber. Free GTP and GDP, which in this case represent >96% of the total nucleotide present at any one time in the reaction mixture, do not give an LD signal because they are not able to be aligned. Hence, LD provides an exquisite method for probing the on-enzyme conformation of nucleotide.

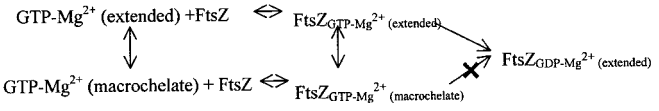
Consequences of Elevated Mg^{2+} on GTPase Activity—It is established that GTPases require the presence of a divalent cation (usually Mg^{2+}) coordinating the phosphates of the GTP for efficient catalysis to proceed. The metal ion usually bridges between the oxygen atoms of the terminal 2 phosphates of the nucleotide and oxygens from bulk solvent or protein side chains. GTPases are also exquisitely sensitive to the type of divalent cation that performs this coordination. The GTPase activity of FtsZ, for example, is completely abolished in the presence of Ca^{2+} instead of Mg^{2+} (Table II). It is, however, often erroneously assumed that metal-chelated GTP in solution only occurs in the extended conformation appropriate for catalysis. In fact, studies by Sigel *et al.* (54) demonstrate that GTP can exist in two very different metal bound conformations, extended and a folded macrochelate. The macrochelate occurs as a result of an interaction between the metal ion and the N7 nitrogen on the guanine, leading to a folded conformation of the nucleotide. This conformation is restricted to purines and is more prevalent in guanine compared with adenine. The interaction between a metal ion (M^{2+}) and GTP can, thus, be written,

$$\text{GTP} + M^{2+} \Rightarrow \text{GTP-}M^{2+}\ \text{(extended)} \overset{K_{eq}}{\Longleftrightarrow} \text{GTP-}M^{2+}\ \text{(macrochelate)}$$

REACTION I

where M^{2+} is a divalent metal ion.

At saturating concentrations of metal ion (dissociation constants for Ca^{2+}, Mg^{2+}, and Cu^{2+} between 0.05 and 2 mM, 0.1 and 2 mM, and 40 nM and 0.25 mM, respectively, depending on the ionization of the GTP), the relative proportions of the nucleotide in each conformation was determined by the equilibrium constant K_{eq}. Sigel *et al.* show (54) that this constant is determined by the type of metal ion present, with the maximum conversion to the macrochelate state occurring in the presence of Cu^{2+}. Solutions of GTP with Ca^{2+}or Mg^{2+}, thus, contain 24 and 21% of the macrochelate form, respectively, at full saturation. The question is now posed as to what implications this has for the on-enzyme population of GTP. It has been shown that the FtsZ GTPase activity increases as Mg^{2+} concentration increases up to an optimal level (26). This is consistent with the formation of a catalytically competent GTP-Mg^{2+} complex on the enzyme (including chelation of the metal by protein side chains). At these elevated concentrations of metal ion, formation of the GTP-Mg^{2+} complex becomes favored in solution until almost all GTP is in complex with Mg^{2+}. For Mg^{2+}, ~20% of this GTP-Mg^{2+} complex is in the form of the macrochelate. When this situation is reached in the presence of FtsZ we have the set of potential steps shown in Reaction II. The upper row of reactions represents the binding of GTP-Mg^{2+} in the extended conformation to FtsZ, which is the productive pathway leading to GTP hydrolysis. The binding of this conformation is analogous to the binding of the catalytically competent conformation observed in crystal structures of GTPases with bound GTP and Mg^{2+}. The lower row of reactions represents an equivalent set of events for the macrochelate form of GTP-Mg^{2+}. Although macrochelate formation has been detected and measured in solution (54), there are no data on its interactions with GTPases. Our data from near UV LD at elevated concentrations of Mg^{2+}, however, indicates the presence of GTP in the active site in a radically altered conformation compared with that seen at lower Mg^{2+} concentrations. We propose that these data obtained with FtsZ provide the first evidence of GTP-Mg^{2+} in the macrochelate conformation (or in a macrochelate-like conformation) bound to a GTPase. Data from structural studies of the GTPase active site suggest that hydrolysis of macrochelate GTP is unlikely to occur because the precise positioning of the phosphates and metal ion in the active site is crucial for catalysis. Our data support this because in each case we detected the signature LD signal for the alternative GTP conformation. At elevated Mg^{2+} concentrations we also detected a decrease in GTPase activity that



REACTION II

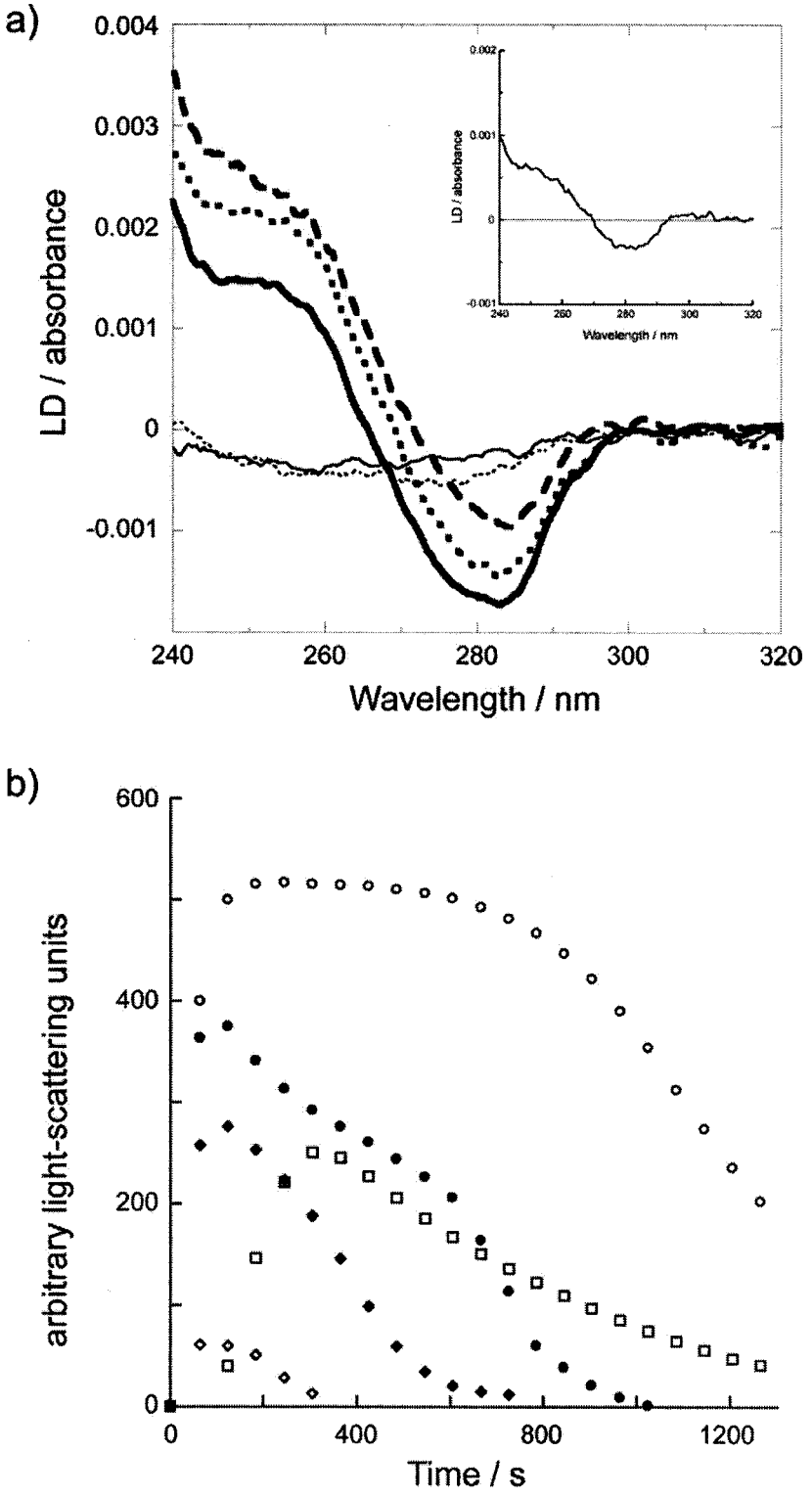


FIG. 6. **Ca^{2+} alone or elevated levels of Mg^{2+} can also induce guanine flipping and FtsZ polymer bundling.** *a*, LD spectra of FtsZ at 11 μM in 50 mM MES buffer, pH 6.5, and 50 mM KCl at 23 °C after the addition of GTP to 0.2 mM in the presence of no divalent cations (*thin line*), 10 mM $MgCl_2$ (*thin-dotted line*), 10 mM $CaCl_2$ (*thick line*), 10 mM $MgCl_2$ + 10 mM $CaCl_2$ (*thick-dotted line*), or 40 mM $MgCl_2$ (*thick-dashed line*). The *inset* shows the spectrum in the presence of 0.1 mM $CuCl_2$. *b*, FtsZ polymerization monitored by light-scattering after the addition of 0.1 mM GTP in the presence of 10 mM $MgCl_2$ (*open diamonds*), 10 mM $MgCl_2$ + 10 mM $CaCl_2$ (*open circles*), 30 mM $MgCl_2$ (*filled diamonds*), 40 mM $MgCl_2$ (*filled circles*), or 60 mM $MgCl_2$ (*open squares*).

matches the onset of bundling. This is a consequence of the changes in the amounts of free Mg^{2+} and free GTP as well as the relative amounts of the two possible GTP-Mg^{2+} complexes. From this it is a logical extension that these in turn are related to the values for the association constants for formation of the GTP-Mg^{2+} complex (both on and off enzyme) as well as the equilibrium constant for the interconversion between the extended and macrochelate form of the Mg^{2+}-GTP complex. Although direct determination of all these constants is beyond the realm of this study, it seems likely that the increase in the absolute concentration of the macrochelate form at high magnesium concentrations (tending to 20% of all GTP) is responsible for the decrease in GTPase. This is particularly likely as the FtsZ-GTP (macrochelate) conformation is likely to be catalytically inactive and can, thus, be treated as a competitive inhibitor of FtsZ.

Consequences of elevated Mg^{2+} and Ca^{2+} on FtsZ Bundling—It has previously been observed that the presence of high concentrations of the catalytically inert metal ion Ca^{2+} can induce the bundling of FtsZ polymers (27, 33), and we

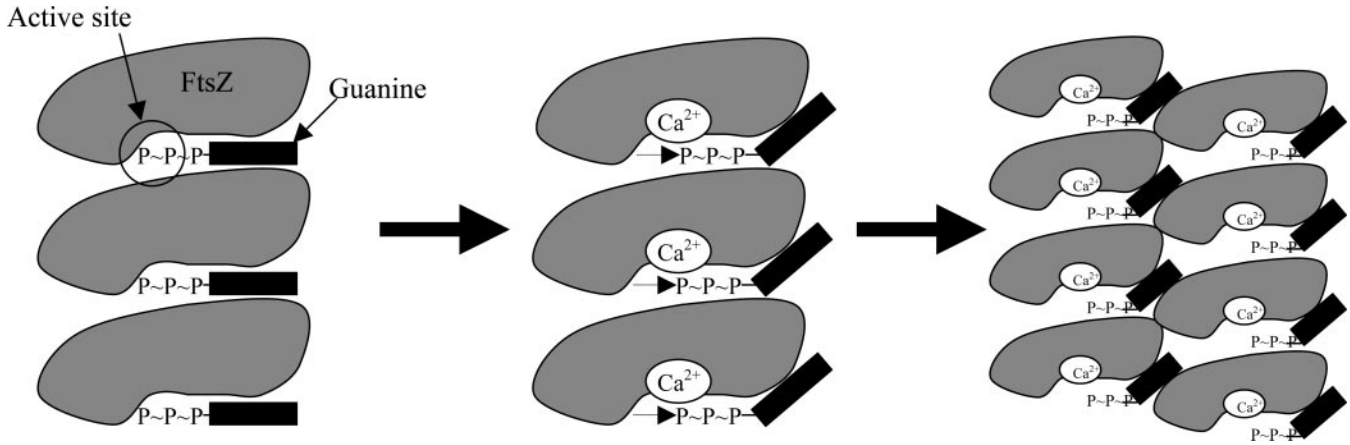


FIG. 7. **Model for Ca^{2+}-induced FtsZ bundling.** Ca^{2+} binds to FtsZ, altering the bound conformation of GTP, leading to displacement of the phosphates from the active site and, hence, a reduction in GTPase activity. Flipping of the guanine moiety into the macrochelate conformation moves it to a more surface-exposed position, where it forms a new interface for bundling.

observe similar bundling at elevated Mg^{2+} concentrations (Fig. 5*e*). In both cases we also observe flipping of the guanine group indicative of GTP bound to FtsZ adopting an alternative conformation. As outlined in the previous section, elevated concentrations of divalent metal cations in solution with GTP lead to the formation of increasing amounts of macrochelate containing metal and nucleotide. We, therefore, suggest that polymer bundling and macrochelate formation may be linked. However, we recognize a number of possible mechanisms by which calcium ions could induce FtsZ polymer bundling; 1) Ca^{2+} acts to buffer negative charges on the surface of FtsZ and, hence, reduces repulsive forces and induces association between polymers (28), 2) Ca^{2+} allows the formation of stable FtsZ polymers by virtue of the ion acting as an analogue of Mg^{2+} but without allowing hydrolysis (the increased persistence of these polymers allows more time for lateral bundling to occur), 3) Ca^{2+} binds to the FtsZ polymer, resulting in a conformational change that aids bundling (the conformational change occurs in either FtsZ or GTP or both).

Our data allow us to distinguish between these possible mechanisms. We observe very little change in the backbone conformation of FtsZ between bundled and non-bundled polymers. This indicates that if a conformation change in the protein architecture is important (mechanism 3) it involves only small changes in backbone torsions and perhaps side chain movements. We observe a signature LD signal for an alternative GTP conformation at concentrations of Ca^{2+} that induce bundling (Fig. 4*d*). We propose that this represents a nucleotide conformation that is quite different to the extended conformations observed in x-ray crystal structures. The question is then whether Ca^{2+} induces the conformational change directly or by binding to FtsZ and in some way alters the conformation of the nucleotide binding site with a resulting "knock on" effect on the cofactor conformation. Because an alternative conformation of GTP is populated in solution at high metal ion concentrations (54), we favor a direct interaction between Ca^{2+} and GTP leading to the observed LD signal. In each case where we have detected the signature LD signal for the alternative GTP conformation, we have also detected polymer bundling. We therefore conclude that the presence of GTP in the alternative (macrochelate-like) conformation is linked to fiber bundling. A second link, between bundling and a decrease in GTPase, still holds and it seems likely that the cause of the reduction of the GTPase is the adoption of the alternative conformation of GTP bound to FtsZ.

This model has allowed us to make and test a prediction. If the alternative GTP conformation is similar to the macrochelate structure and if the formation of this structure is important in bundling, then an agent that induces the macrochelate conformation more efficiently than Ca^{2+} or Mg^{2+} would be a more potent bundler of FtsZ polymers. Cu^{2+} is the most potent macrochelate-inducing metal ion so far studied (>90% of GTP-Cu^{2+} is in the macrochelate conformation in solution) (54). We, therefore, tested the effect of adding $CuCl_2$ to standard FtsZ polymerization reactions. Our results showed that Cu^{2+} induces the formation of FtsZ polymer bundles at much lower concentrations (down to 0.1 mM) than Ca^{2+} (Fig. 5*f*). Incubation of FtsZ with $CuCl_2$ at higher concentrations (10 mM) resulted in extensive protein aggregation as reported previously for copper and the other divalent cations Mn^{2+}, Fe^{2+}, Co^{2+}, and Ni^{2+} (27). It is significant that Cu^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, and Ni^{2+} all have more than three times the preponderance for forming the macrochelate than do either Ca^{2+} or Mg^{2+} (54). Furthermore, we find that FtsZ polymers formed in the presence of Cu^{2+} show the alternative GTP conformation signature LD signal at a lower concentration (0.1 mM) than either Mg^{2+} or Ca^{2+} (Fig. 6*a*, *inset*). This provides strong support for the proposal that formation of a macrochelate conformation of GTP bound to FtsZ leads to fiber bundling. It may also be considered that the effect of Cu^{2+} may be due to an interaction between the $Cu^{2+}$and cysteine residues in the protein (a well characterized interaction in other systems). However, examination of the primary sequence of *E. coli* FtsZ shows the lack of any cysteines.

It should be noted that if bundling occurs due to the effect of cationic buffering, it would be expected that low concentrations of Cu^{2+} would induce little or no bundling. Because we find the converse to be true, we can rule out mechanism 1 as the sole driving force behind FtsZ polymer bundling.

Taken together these results indicate that FtsZ bundling is inherently linked to a conformational change in the GTP. The conditions under which this conformation is formed match closely to those that favor formation of the macrochelate structure of GTP-Me^{2+} in solution. The result of this conformational change in polymers formed using Mg^{2+} is to produce polymers with reduced GTP turnover that are, hence, less dynamic. This could be said to be consistent with mechanism 2 proposed above. However, since a number of situations where reduced GTPase activity does not lead to polymer bundling have been reported (26, 33, 55, 56), a reduction in dynamics must not in itself be responsible for bundle formation (conversely, it is interesting to note that bundling may always result in reduced dynamics, since all studies of FtsZ bundling where GTPase activity was tested report that activity decreases (23, 27, 33, 57,

58).) If FtsZ bundling is thought to be solely the result of a reduced GTPase activity, then it would also be expected that at saturation the equally catalytically inert cations Cu^{2+} and Ca^{2+} should have very similar effects on bundling. As we have seen, this is very much not the case. Thus, the increased bundling potency of the Cu^{2+} over the equally catalytically inert Ca^{2+} can be more easily explained by a secondary, catalytically unrelated effect of the alternative GTP conformation. Examination of the x-ray crystal structure of GDP-FtsZ shows the guanine moiety to be close to the surface of the protein. Therefore, we propose a model whereby the alternative on-enzyme conformation of GTP-M^{2+} creates a site for lateral fiber association near the GTP binding pocket (Fig. 7).

Acknowledgment—We thank Dr. Löwe (Cambridge) for the use of the FtsZ protofiber model.

REFERENCES

1. Margolin, W. (2000) *FEMS Microbiol. Rev.* **24,** 531–548
2. Lutkenhaus, J., and Addinall, S. G. (1997) *Annu. Rev. Biochem.* **66,** 93–116
3. Addinall, S. G., and Holland, B. (2002) *J. Mol. Biol.* **318,** 219–236
4. Chen, J. C., and Beckwith, J. (2001) *Mol. Microbiol.* **42,** 395–413
5. Hale, C. A., and de Boer, P. A. J. (2002) *J. Bacteriol.* **184,** 2552–2556
6. Ma, X., Ehrhardt, D. W., and Margolin, W. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 12998–13003
7. Addinall, S. G., and Lutkenhaus, J. (1996) *J. Bacteriol.* **178,** 7167–7172
8. Vaughan, S., Wickstead, B., Gull, K., and Addinall, S. G. (2004) *J. Mol. Evol.* **58,** 19–29
9. Glockner, F. O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Beck, A., Borzym, K., Heitmann, K., Rabus, R., Schlesner, H., Amann, R., and Reinhardt, R. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 8298–8303
10. Miyagishima, S.-Y., Nishida, K., and Kuroiwa, T. (2003) *Trends Plant Sci.* **8,** 432–438
11. Osteryoung, K. W., and Nunnari, J. (2003) *Science* **302,** 1698–1704
12. Gilson, P. R., Yu, X.-C., Hereld, D., Barth, C., Savage, A., Kiefel, B. R., Lay, S., Fisher, P. R., Margolin, W., and Beech, P. L. (2003) *Eukaryot. Cell* **2,** 1315–1326
13. Romberg, L., and Mitchison, T. J. (2004) *Biochemistry* **43,** 282–288
14. Romberg, L., and Levin, P. A. (2003) *Annu. Rev. Microbiol.* **57,** 125–154
15. Erickson, H. P. (1998) *Trends Cell Biol.* **8,** 133–137
16. Scheffers, D. J., de Wit, J. G., den Blaauwen, T., and Driessen, A. J. (2002) *Biochemistry* **41,** 521–529
17. Lowe, J., and Amos, L. A. (1998) *Nature* **391,** 203–206
18. Lowe, J., and Amos, L. A. (1999) *EMBO J.* **18,** 2364–2371
19. Diaz, J. F., Kralicek, A., Mingorance, J., Palacios, J. M., Vicente, M., and Andreu, J. M. (2001) *J. Biol. Chem.* **276,** 17307–17315
20. Lu, C., Reedy, M., and Erickson, H. P. (2000) *J. Bacteriol.* **182,** 164–170
21. Romberg, L., Simon, M., and Erickson, H. P. (2001) *J. Biol. Chem.* **276,** 11743–11753
22. Lowe, J., and Amos, L. A. (2000) *Biol. Chem.* **381,** 993–999
23. Gueiros-Filho, F. J., and Losick, R. (2002) *Genes Dev.* **16,** 2544–2556
24. Oliva, M. A., Huecas, S., Palacios, J. M., Martin-Benito, J., Valpuesta, J. M., and Andreu, J. M. (2003) *J. Biol. Chem.* **278,** 33562–33570
25. Erickson, H. P., Taylor, D. W., Taylor, K. A., and Bramhill, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 519–523
26. Mukherjee, A., and Lutkenhaus, J. (1998) *EMBO J.* **17,** 462–469
27. Yu, X. C., and Margolin, W. (1997) *EMBO J.* **16,** 5455–5463
28. Hale, C. A., Rhee, A. C., and de Boer, P. A. (2000) *J. Bacteriol.* **182,** 5153–5166
29. Gonzalez, J. M., Jimenez, M., Velez, M., Mingorance, J., Andreu, J. M., Vicente, M., and Rivas, G. (2003) *J. Biol. Chem.* **278,** 37664–37671
30. Addinall, S. G., Cao, C., and Lutkenhaus, J. (1997) *J. Bacteriol.* **179,** 4277–4284
31. Stricker, J., Maddox, P., Salmon, E. D., and Erickson, H. P. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 3171–3175
32. Trusca, D., Scott, S., Thompson, C., and Bramhill, D. (1998) *J. Bacteriol.* **180,** 3946–3953
33. Mukherjee, A., and Lutkenhaus, J. (1999) *J. Bacteriol.* **181,** 823–832
34. Scheffers, D. J., den Blaauwen, T., and Driessen, A. J. (2000) *Mol. Microbiol.* **35,** 1211–1219
35. Rivas, G., Fernandez, J. A., and Minton, A. P. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 3150–3155
36. Small, E., and Addinall, S. G. (2003) *Microbiology* **149,** 2235–2242
37. Mukherjee, A., Cao, C., and Lutkenhaus, J. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2885–2890
38. Sivasothy, P., Dafforn, T. R., Gettins, P. G. W., and Lomas, D. A. (2000) *J. Biol. Chem.* **275,** 33663–33668
39. Rodger, A., Rajendra, J., Marrington, R., Ardhammar, M., Nordén, B., Hirst, J. B., Gilbert, A. T. B., Dafforn, T. R., Halsall, D. J., Woolhead, C. A., Robinson, C., Pinheiro, T. J., J., K., Seymour, M., Perez, N., and Hannon, M. J. (2002) *Phys. Chem. Chem. Phys.* **4,** 4051–4057
40. Rodger, A., Rajendra, J., Marrington, R., Mortimer, R., Andrews, T., Hirst, J. B., Gilbert, A. T. B., Halsall, D. J., Dafforn, T. R., Ardhammar, M., Nordén, B., Woolhead, C. A., Robinson, C., Pinheiro, T., J., K., Seymour, M., Perez, N., and Hannon, M. J. (2002) in *Biophysical Chemistry: Membranes and Proteins* (Templer, R. H., ed) pp. 3–19, Royal Society of Chemistry, Cambridge, UK
41. Dafforn, T. R., Rajendra, J., Halsall, D. J., Serpell, L. C., and Rodger, A. (2004) *Biophys. J.* **86,** 404–410
42. Marrington, R., Dafforn, T. R., Halsall, D. J., Andrews, T., and Rodger, A. (2004) *Biophys. J.* **87,** 2002–2012
43. Mukherjee, A., and Lutkenhaus, J. (1998) *Methods Enzymol.* **298,** 296–305
44. Lu, C., and Erickson, H. P. (1998) *Methods Enzymol.* **298,** 305–313
45. Pandya, M. J., Spooner, G. M., Sunde, M., Thorpe, J. R., Rodger, A., and Woolfson, D. N. (2000) *Biochemistry* **39,** 8728–8734
46. Nordh, J., Deinum, J., and Norden, B. (1986) *Eur. Biophys. J.* **14,** 113–122
47. Lockhart, A., and Kendrick-Jones, J. (1998) *FEBS Lett.* **430,** 278–282
48. Sossong, T. M., Jr., Brigham-Burke, M. R., Hensley, P., and Pearce, K. H., Jr. (1999) *Biochemistry* **38,** 14843–14850
49. Miki, M., and Mihashi, K. (1976) *Biophys. Chem.* **6,** 101–106
50. Lu, C., Stricker, J., and Erickson, H. P. (1998) *Cell Motil. Cytoskeleton* **40,** 71–86
51. Clark, L. B. (1977) *J. Am. Chem. Soc.* **99,** 3934–3938
52. Zaloudek, F., Novros, J. S., and Clark, L. B. (1984) *J. Am. Chem. Soc.* **107,** 7344–7351
53. Rodger, A., and Nordén, B. (1987) *Circular and Linear Dichroism*, Oxford University Press, Oxford
54. Sigel, H., Bianchi, E. M., Corfu, N. A., Kinjo, Y., Tribolet, R., and Martin, R. B. (2001) *Chemistry* **7,** 3729–3737
55. Lu, C., Stricker, J., and Erickson, H. P. (2001) *BMC Microbiol.* **1,** 7
56. White, E. L., Ross, L. J., Reynolds, R. C., Seitz, L. E., Moore, G. D., and Borhani, D. W. (2000) *J. Bacteriol.* **182,** 4028–4034
57. Beuria, T. K., Krishnakumar, S. S., Sahar, S., Singh, N., Gupta, K., Meshram, M., and Panda, D. (2003) *J. Biol. Chem.* **278,** 3735–3741
58. Santra, M. K., Beuria, T. K., Banerjee, A., and Panda, D. (2004) *J. Biol. Chem.* **279,** 25959–25965
59. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723
60. Johnson, W. C., Guex, N., and Peitsch, M. C. (1999) *Proteins* **35,** 307–312

ELSEVIER

Linear dichroism of biomolecules: which way is up?

Timothy R Dafforn[1] and Alison Rodger[2]

Understanding the organization of molecules in naturally occurring ordered arrays (e.g. membranes, protein fibres and DNA strands) is of great importance to understanding biological function. Unfortunately, few biophysical techniques provide detailed structural information on these non-crystalline systems. UV, visible and IR linear dichroism have the potential to provide such information. Recent advances in technology and simulations allow this potential to be fulfilled, and can now provide a detailed understanding of the molecular mechanisms of such fundamental biological processes as amyloid fibre formation and membrane protein folding.

Addresses
[1] Department of Biosciences, University of Birmingham, Birmingham B15 2TT, UK
e-mail: T.R.Dafforn@bham.ac.uk
[2] Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK
e-mail: A.Rodger@warwick.ac.uk

Current Opinion in Structural Biology 2004, **14**:541–546

This review comes from a themed issue on
Biophysical methods
Edited by Arthur G Palmer III and Randy J Read

Available online 15th September 2004

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2004.08.005

Abbreviations

ATR	attenuated total reflectance
CD	circular dichroism
FT	Fourier transform
LD	linear dichroism

Introduction

The Protein Data Bank (PDB) currently contains the coordinates of some 25000 biomolecules and is continuing to grow at a near exponential rate. It is, however, known that reliance on X-ray crystallography, NMR and electron microscopy for atomic-resolution data has meant that the database is highly skewed towards soluble, globular proteins, a subset that makes up only around 30% of gene products. The 'hidden' molecules can be classed as membrane proteins, fibrous proteins and unstructured proteins. In many cases, directly elucidating the atomic structure of these proteins has proved impossible and, instead, our quest for knowledge proceeds by a slower more circuitous route. Such studies inevitably involve the use of a range of biophysical techniques, many based around molecular spectroscopy, combined with often clever and elegant biochemical analysis. The end result is a 'collage' of information that is used to form an overall impression of the structure of the protein. In this review, we will examine one of the emerging spectroscopic techniques that is providing new information on the structures of membrane and fibrous proteins — linear dichroism.

Linear dichroism

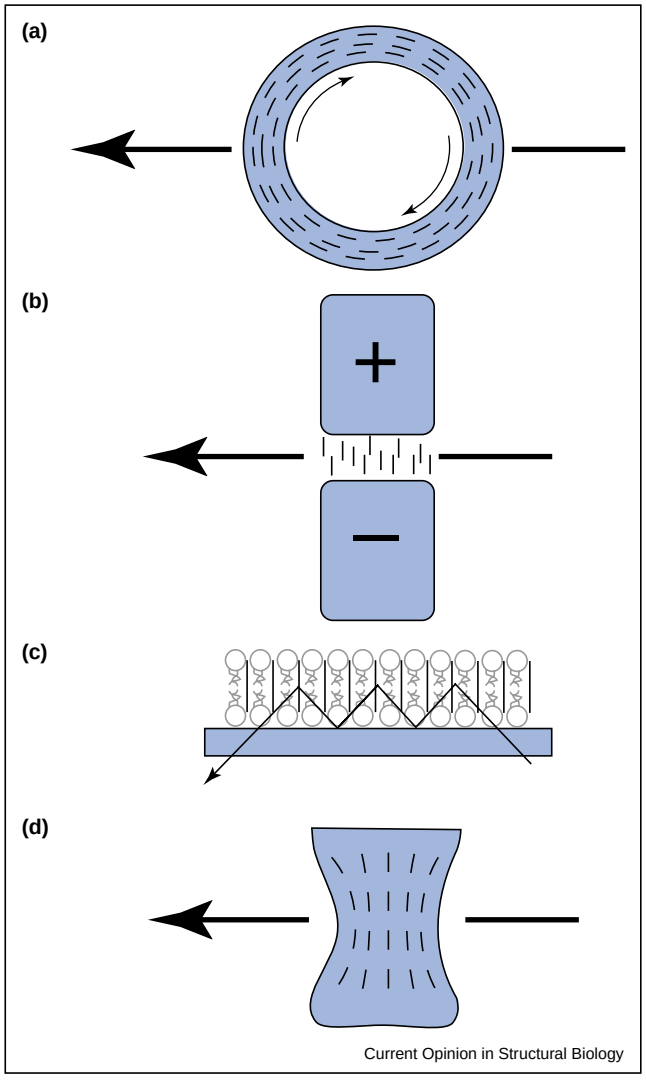
In its most basic form, the absorption of different wavelengths of light by biomolecules has been used to gain information on their chemical make-up. Light, however, has another quality that can be used to discover things about the structure of biomolecules, namely polarisation. The most commonly used polarised light is circularly polarised; its importance in biophysics stems from the fact that the majority of the constituents of biomolecules and the conformations they adopt are chiral. These chiral structures interact with the left and right circularly polarised light to differing extents. The absorbance differences give rise to a circular dichroism (CD) spectrum that can be deconvoluted to give information on the conformation of biomacromolecules. Linearly polarised light (usually horizontally or vertically polarised) gives rise to a technically related method, linear dichroism (LD), a measure of the difference in absorption of the horizontally and vertically polarised light beams by a sample that is aligned preferentially along the horizontal direction. LD has been somewhat the poor cousin to CD in biophysical chemistry in the past, the over-riding reason being that, to obtain an LD signal from a sample, the molecules within that sample have to be regularly aligned parallel to one of the polarisation directions of the light.

Inducing molecular alignment

Molecular alignment ensures that the electronic transitions that would normally interact with the incident light are aligned and hence can coherently interact with the polarised light. Indeed, a randomly oriented solution of biomolecules shows no LD signal at all. The need for alignment has severely restricted the use of LD in biology on two counts. Firstly, the molecules must have some kind of asymmetric characteristic, such as a charge dipole or elongated shape to be aligned. Secondly, apparatus to achieve sufficient alignment while preserving sample integrity has not been readily available. Common alignment methods include: shear flow orientation [1–3]; electromagnetic orientation [4]; thin film orientation [5,6]; and gel squeezing [7–10].

It is apparent that these different methods (Figure 1) are appropriate for different types of biomolecules. Shear flow relies on the use of molecules with high aspect ratios;

Figure 1



The four techniques most commonly used to induce molecular alignment in biomolecules during LD experiments. **(a)** Flow alignment of a solution of high aspect ratio biomolecules by a shear flow gradient produced by the spinning of a cylinder within a concentric sleeve. **(b)** Magnetic alignment induced by the presence of an electromagnetic field. **(c)** Thin film alignment induced by assembly of a biomolecule into a lipid bilayer on a crystal used in an IR ATR experiment. **(d)** Squeeze alignment induced by deformation of a biomolecule-loaded gel.

electromagnetic alignment requires significant charge dipole properties.

Linear dichroism and DNA binding

Most literature reports of the application of LD to structural studies of biomacromolecules involve DNA and DNA–ligand systems. This derives from the interests of most of the laboratories with facilities to build their own Couette flow cells to orient the samples and the ready availability of DNA samples that could be flow oriented. A summary of DNA LD is helpful for understanding its application to other biomolecules.

DNA is composed of aromatic molecules (bases) stacked vertically and twisted about a helix axis. The accessible DNA transitions are all π–π^* transitions of these bases (see Figure 2), whose absorption starts from about 300 nm and runs down into the far UV. LD is often used to probe the orientation of extrinsic chromophores (including DNA-binding drugs). From such studies, it can be shown that, when ligands bind to DNA, they may:

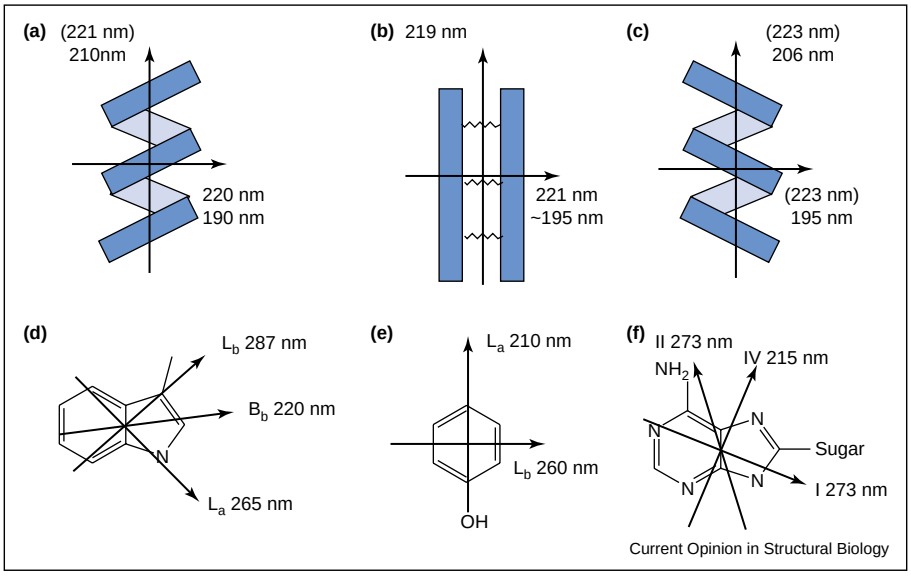
1. Intercalate between the DNA bases, in which case their π–π^* transitions are parallel to those of the DNA bases and their LD signals are the same as those of the local DNA bases [11,12] (e.g. ethidium bromide, Figure 3).
2. Bind along the minor groove, in which case their long axis transitions are $\sim 45^\circ$ from the helix axis, with positive LD signals, and their short axis transitions lie parallel to the DNA bases, with negative LD signals (e.g. 4′,6-diamidino-2-phenylindole [DAPI], Figure 3) [4].
3. Bind in the major groove, where geometry constraints are less rigid [13,14].
4. Associate non-specifically with the DNA backbone, in which case the LD signal will be very small or zero.

LD can also provide detailed information on the effect of such ligands on DNA conformation. For instance, an intercalating ligand (e.g. ethidium bromide, Figure 3) locally stiffens the DNA and thus increases the local orientation; this results in a negative ligand LD of larger magnitude than that of the average DNA base. Another, more extreme, case of DNA perturbation is provided by dimetallo helicates, which intramolecularly coil DNA; in this case, LD is an effective measure of DNA bending rather than giving the orientation of the ligand on the DNA (Figure 4).

Linear dichroism and protein fibres

Studies of protein fibres usually rely heavily on signals that provide little or no direct structural information (e.g. light scattering, intrinsic protein fluorescence and extrinsic fluorescence of amyloid-specific dyes) and allow examination of only the kinetics of formation. At best, CD and FTIR (Fourier transform IR) absorbance data have been used to provide some information on the secondary structure components of fibres. Studies of protein fibres that are polymers of monomeric globular proteins (e.g. F-actin, tubulin microtubules) have also used such optical techniques. However, in these cases, high-resolution data from X-ray crystal or NMR structures of the monomeric units that make up the fibre help structural characterisation. The orientation of the protein monomers and associated chromophores (ATP, GTP, etc.) within the fibre is, however, still hard to ascertain. LD provides a useful

Figure 2

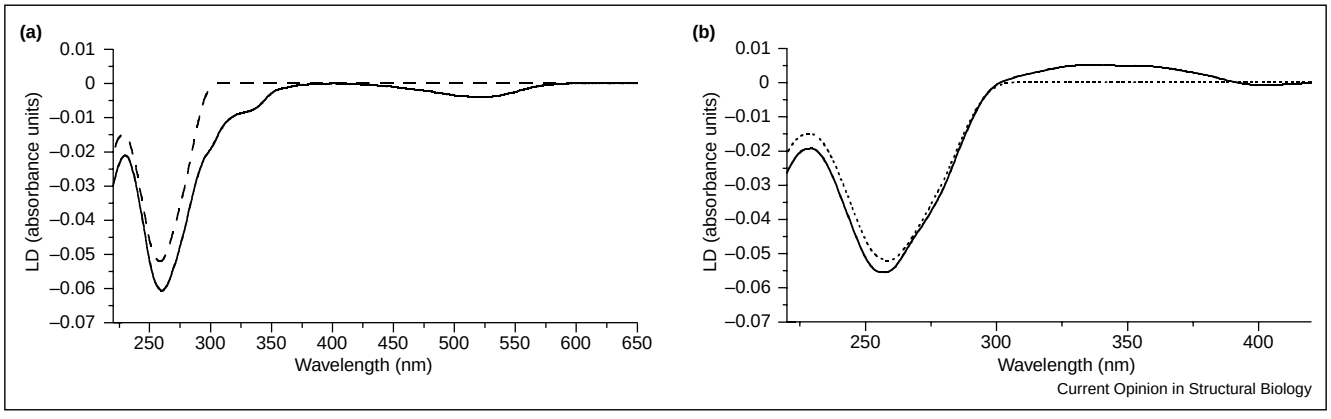


The orientations of the various UV polarisation moments in **(a)** α helix, **(b)** β sheet, **(c)** polyproline type II helix, **(d)** tryptophan, **(e)** tyrosine and **(f)** adenosine chromophores. Arrows indicate the transition moment polarisations and the approximate wavelength maxima of the transitions are given near the arrows. In some cases, the common notation for the transition is also given. Wavelengths in parentheses indicate that the intensity of this transition is weak.

extra tool for the study of these proteins. In early work on the LD spectra of F-actin [15,16], it was demonstrated that LD could provide information on the orientation of both secondary structure elements, aromatic amino acids and bound nucleotide within F-actin. In these studies [15,16], shear flow alignment was used to order F-actin in solution, pointing the way forward for other such studies. A more recent study of nitrosylated haemoglobin S has used LD to examine the polymerisation of these mutant proteins [17]. Unfortunately, the limited availability of flow alignment systems (as discussed above) has meant that, until recently, this work could not be extended to other systems. However, with the development of low volume/high alignment Couette systems in our own laboratories, progress is again being made [18,19•].

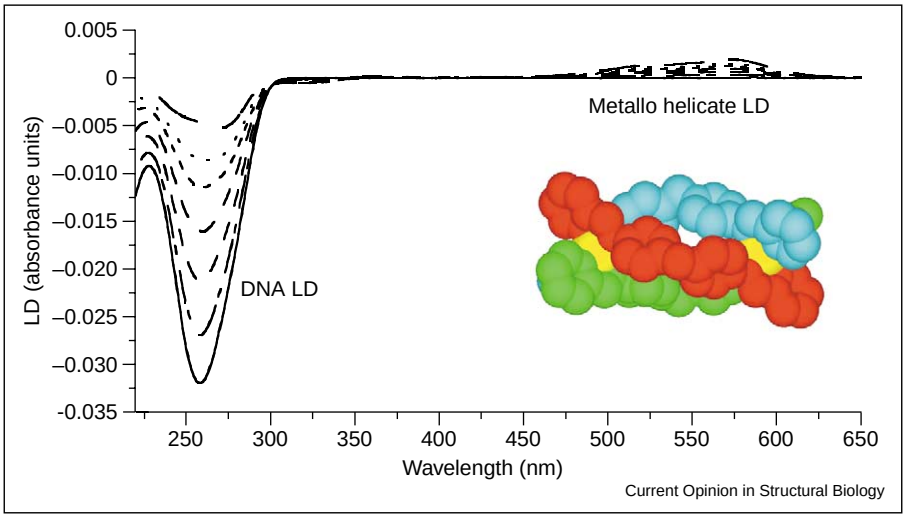
Our first study of a fibrous protein was an examination of artificial self-associating coiled-coil fibres [20]. These assemblies represent one of the simplest forms of protein fibre (containing only α helix and no aromatic residues).

Figure 3



LD of DNA and DNA–ligand systems. **(a)** LD of calf thymus DNA (1000 μM base, dashed line) and the DNA plus an ethidium bromide intercalator (50 μM, solid line). **(b)** LD of calf thymus DNA (1000 μM base, dashed line) and the DNA plus a minor groove binder (diaminophenyl indole, 50 μM, solid line) [32].

Figure 4



Reduction in the DNA (500 μm) LD signal (at ~260 nm) upon addition of a DNA-bending dimetallo helicate (inset) (whose LD is observed at ~550 nm). The helicate concentration increases with decreasing magnitude of DNA LD signal: 0 μM, 10 μM, 20 μM, 30 μM, 50 μM, 60 μM, 100 μM.

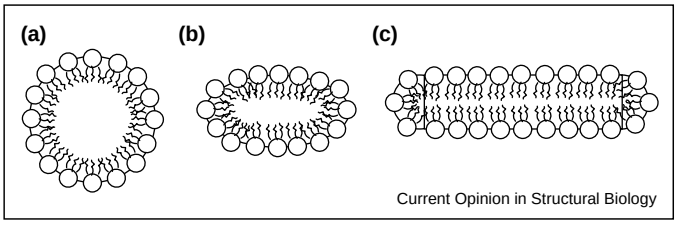
The experimental data allowed determination of the directions of the electronic transitions for the α helix. Perhaps more importantly, it proved that the alignment system was usable with protein samples and was able to provide information on the orientation of the α helix with respect to the fibre axis. This success initiated a further study [19•], whereby a range of fibres were studied with the aim of examining the signals from each specific type of secondary structure. The fibres studied included the all-β-sheet amyloid fibre formed by the Alzheimer's peptide, the polyproline type II fibre formed by collagen, the α/β fibre formed by α_1-antitrypsin and F-actin. Examination of the results from these studies has allowed the positive assignment of the features that correlate with the various secondary structures present within the fibres. Another important by-product of this study was the observation that, despite the shear force that the sample is experiencing, in all cases no indication of fragmentation of the sample was observed. The potential application of LD for elucidating the kinetics of fibre formation was also highlighted during this work, the polymerisation process from actin monomer to fully formed polymer being easily monitored. The results from these studies demonstrate clearly the potential importance of LD in the study of protein fibre structure.

Linear dichroism and membrane proteins

Membrane proteins present a specific set of difficulties to biochemists and structural biologists, their high insolubility and requirement for a membrane environment making them hard to handle and examine. LD (both IR and visible-UV) can provide detailed information on the membrane-bound forms of these proteins that is difficult to obtain by other methods. With membrane proteins, the requirement for insertion into the near two-dimensional plane of the membrane helps to provide the alignment required for LD experiments because, if you can orient the membrane, you can orient the protein.

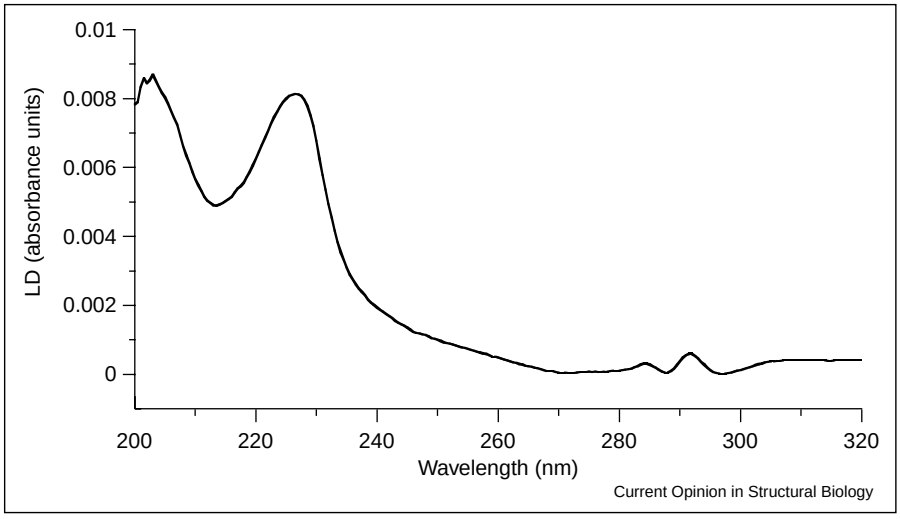
Most IR LD has involved attenuated total reflectance (ATR), whereby the light beam travels within a crystal upon which the sample — in this case, usually a lipid

Figure 5



Schematic illustration of **(a)** a liposome, **(b)** a shear-deformed liposome and **(c)** a model system with an orientation equivalent to a shear-deformed liposome.

Figure 6



LD spectra of gramicidin (head-to-tail dimer, 0.5 mg in 2.5 mg/mL lipid in phosphate buffer) in soybean liposomes prepared by adding ethanol-solubilised gramicidin to an aqueous extruded liposome solution. A non-rotating sample baseline was subtracted [2].

bilayer with embedded protein — is immobilised [5]. The sample is often dehydrated to improve alignment. IR studies focus on the amide I and amide II bands, as these provide information on the secondary structure elements present. For example, Bechinger *et al.* [6] have performed a systematic study of four membrane-resident biomolecules using IR ATR and have demonstrated that the technique can be used to directly determine the angle of inclination of α helices with respect to the membrane. Such studies have been further refined by isotope labelling experiments [21,22•,23] and improved theoretical analysis [24••].

For work in the visible region (e.g. examining the orientation of membrane-bound prosthetic groups [25]), two alignment systems are used: flow and squeezed gel. The use of visible LD has proved particularly useful in the study of membrane proteins containing haem groups (e.g. light-harvesting complexes [26,27] and cytochromes). These studies (usually carried out at low temperature [27,28]) provide detailed information on the orientation of the chromophores within the protein and hence with respect to each other.

Using flow alignment, light scattering from membranes is considerable and cannot be ignored. Attempts to minimise scattering include using micro-path length (<50 μm) Couettes [18], refractive index matching solvents [29] or a focusing lens to collect the scattered light for detection [19•]. The key to flow alignment of membrane proteins was the discovery [30] that small molecules could be flow oriented in liposomes in a Couette flow cell because the flow distorts the liposomes into an elliptical shape that is sufficiently elongated to provide an orientation axis. A range of membrane-bound proteins have been shown to give LD spectra using this method. β Sheets show little signal until the 200 nm transition, as the opposing polarisations (Figure 2) of the 220 nm n–π* transition lead to little net LD. α Helices, by way of contrast, show LD signals for all transitions. When they are inserted in the membrane parallel to the lipids (and thus perpendicular to the long axis of the flow-distorted liposome), the 220 nm region is positive, the 208 nm region negative and the 200 nm region positive, with the signal being given by:

$$LD^r = \frac{3S}{4}\left(1 - 3\cos^2\alpha_i\right) \quad (1)$$

where α_i is the angle the transition moment of interest makes with the normal to the cylinder surface (i.e. to the lipid long axis; Figure 5). The LD spectrum of the membrane-spanning helical protein gramicidin is shown in Figure 6. The large number of tryptophans in this protein means that the aromatic chromophore signal at 280 nm is visible on the same intensity scale as the backbone signal (from 240 nm downwards). The 210 nm negative signal appears as a positive minimum due to overlapping positive signals either side of the signal observed. More recently, the orientation of the peptide penetratin on artificial liposomes has been analysed by LD [31].

Conclusions

LD is one of the rare biophysical techniques to provide information on the orientations of biomolecules in ordered arrays. Recent advances in technical and theoretical aspects of LD now mean that its application to large areas of biology is imminent.

Acknowledgements

TR Dafforn is an MRC Career Development Fellow.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Rodger A: **Linear dichroism**. *Methods Enzymol* 1993, **226**:232-258.

2. Rodger A, Rajendra J, Marrington R, Ardhammar M, Nordén B, Hirst JD, Gilbert ATB, Dafforn TR, Halsall DJ, Woolhead CA: **Flow oriented linear dichroism to probe protein orientation in membrane environments**. *Phys Chem Chem Phys* 2002, **4**:4051-4057.

3. Norden B, Kubista M, Kurucsev T: **Linear dichroism spectroscopy of nucleic acids**. *Q Rev Biophys* 1992, **25**:51-170.

4. Nguyen B, Hamelburg D, Bailly C, Colson P, Stanek J, Brun R, Neidle S, Wilson WD: **Characterisation of a novel DNA minor-groove complex**. *Biophys J* 2004, **86**:1028-1041.

5. Grimard V, Vigano C, Margolles A, Wattiez R, van Veen W, Konings WN, Ruysschaert J-M, Goormaghtigh E: **Structure and dynamics of the membrane-embedded domain of LmrA investigated by coupling polarised ATR-FTIR spectroscopy and $^{1}H/^{2}H$ exchange**. *Biochemistry* 2001, **40**:11876-11886.

6. Bechinger B, Ruysschaert J-M, Goormaghtigh E: **Membrane helix orientation from linear dichroism of infrared attenuated total reflection spectra**. *Biophys J* 1999, **76**:552-563.

7. Hemelrijk PW, Kwa SLS, van Grondelle R, Dekker JP: **Spectroscopic properties of LHC-II, the main light-harvesting chlorophyll a/b protein complex from chloroplast membranes**. *Biochim Biophys Acta* 1992, **1098**:156-166.

8. Tapie P, Haworth P, Hervo G, Breton J: **Orientation of the pigments in the thylakoid membrane and in the isolated chlorophyll-protein complexes of higher plants. III. A quantitative comparison of the low-temperature linear dichroism spectra of thylakoids and isolated pigment-protein complexes**. *Biochim Biophys Acta* 1982, **682**:339-344.

9. Haworth P, Tapie P, Arntzen C, Breton J: **Orientation of pigments in the thylakoid membrane and in the isolated chlorophyll-protein complexes of higher plants. II. Linear dichroism spectra of isolated pigment-protein complexes oriented in polyacrylamide gels at 300 and 100 K**. *Biochim Biophys Acta* 1982, **682**:152-159.

10. Abdourakhmanov I, Ganago AO, Erokhin YE, Solovev A, Chugunov V: **Orientation and linear dichroism of the reaction centers from *Rhodopseudomonas sphaeroides* R-26**. *Biochim Biophys Acta* 1979, **546**:183-186.

11. Dalla Via L, Gia O, Marciani Magno S, Da Settimo A, Primofiore G, Da Settimo F, Simorini F, Marini AM: **Dialkylaminoalkylindolonaphthyridines as potential antitumour agents: synthesis, cytotoxicity and DNA binding properties**. *Eur J Med Chem* 2002, **37**:475-486.

12. Patel KK, Plummer EA, Darwish M, Rodger A, Hannon MJ: **Aryl substituted ruthenium bis-terpyridine complexes: intercalation and groove binding with DNA**. *J Inorg Biochem* 2002, **91**:220-229.

13. Lee S, Lee YA, Lee HM, Lee JY, Kim DH, Kim SK: **Rotation of periphery methylpyridine of meso-tetrakis(n-N-methylpyridiniumyl)porphyrin (n = 2, 3, 4) and its selective binding to native and synthetic DNAs**. *Biophys J* 2002, **83**:371-381.

14. Moldrheim E, Hannon MJ, Meistermann I, Rodger A, Sletten E: **Interaction between a DNA oligonucleotide and a dinuclear iron(II) supramolecular cylinder; an NMR and molecular dynamics study**. *J Biol Inorg Chem* 2002, **7**:770-780.

15. Miki M, Mihashi K: **Fluorescence and flow dichroism of F-actin-epsilon-ADP; the orientation of the amine plane relative to the long axis of F-actin**. *Biophys Chem* 1976, **6**:101-106.

16. Higashi S, Kasai M, Oosawa F, Wada A: **Ultraviolet dichroism of F-actin oriented by flow**. *J Mol Biol* 1963, **7**:421-430.

17. Xu X, Lockamy VL, Chen K, Huang Z, Shields H, King SB, Ballas SK, Nichols JS, Gladwin MT, Noguchi CT *et al.*: **Effects of iron nitrosylation on sickle cell hemoglobin solubility**. *J Biol Chem* 2002, **277**:36787-36792.

18. Rodger A, Rajendra J, Marrington R, Mortimer R, Andrews T, Hirst JB, Gilbert ATB, Marrington R, Halsall D, Dafforn TR *et al.*: **Flow oriented linear dichroism to probe protein orientation in membrane environments**. In *Biophysical Chemistry: Membranes and Proteins.* Edited by Templer RH, Leatherbarrow RJ. Cambridge, UK: Royal Society of Chemistry; 2002:3-19.

19. Dafforn TR, Rajendra J, Halsall DJ, Serpell LC, Rodger A:
• **Protein fibre linear dichroism for structure determination and kinetics in a low volume, low wavelength Couette flow cell**. *Biophys J* 2004, **86**:404-410.
The new design of a short path-length LD Couette, combined with a range of protein fibres, provides the first atlas of protein structure spectra for LD.

20. Pandya MJ, Sunde M, Thorne JR, Rodger A, Woolfson DN: **Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis**. *Biochemistry* 2000, **39**:8728-8734.

21. Torres J, Kukol A, Arkin IT: **Use of a single glycine residue to determine the tilt and orientation of a transmembrane helix. A new structural label for infrared spectroscopy**. *Biophys J* 2000, **79**:3139-3143.

22. Kukol A, Torres J, Arkin IT: **A structure for the trimeric MHC**
• **class II-associated invariant chain transmembrane domain**. *J Mol Biol* 2002, **320**:1109-1117.
A good example of the use of FTIR ATR and molecular dynamics to provide information on membrane helix orientation.

23. Torres J, Kukol A, Goodman JM, Arkin IT: **Site-specific examination of secondary structure and orientation determination in membrane proteins: the peptidic (13)C=(18)O group as a novel infrared probe**. *Biopolymers* 2001, **59**:396-401.

24. Marsh D: **Infrared dichroism of isotope-edited alpha-helices**
•• **and beta-sheets**. *J Mol Biol* 2004, **338**:353-367.
An excellent and comprehensive paper on the theory of FTIR LD.

25. Schoepp B, Chabaud E, Breyton C, Vermeglio A, Popot JL: **On the spatial organization of hemes and chlorophyll in cytochrome *b*(6)f. A linear and circular dichroism study**. *J Biol Chem* 2000, **275**:5275-5283.

26. Byrdin M, Jordan P, Krauss N, Fromme P, Stehlik D, Schlodder E: **Light harvesting in photosystem I: modeling based on the 2.5-A structure of photosystem I from *Synechococcus elongatus***. *Biophys J* 2002, **83**:433-457.

27. De Weerd FL, Palacios MA, Andrizhiyevskaya EG, Dekker JP, Van Grondelle R: **Identifying the lowest electronic states of the chlorophylls in the CP47 core antenna protein of photosystem II**. *Biochemistry* 2002, **41**:15224-15233.

28. Croce R, Canino G, Ros F, Bassi R: **Chromophore organization in the higher-plant photosystem II antenna protein CP26**. *Biochemistry* 2002, **41**:7334-7343.

29. Ardhammar M, Lincoln P, Norden B: **Invisible liposomes: refractive index matching with sucrose enables flow dichroism assessment of peptide orientation in lipid vesicle membrane**. *Proc Natl Acad Sci USA* 2002, **99**:15313-15317.

30. Ardhammar M, Norden B, Nielsen PE, Malmstrom BG, Wittung-Stafshede P: ***In vitro* membrane penetration of modified peptide nucleic acid (PNA)**. *J Biomol Struct Dyn* 1999, **17**:33-40.

31. Brattwall CE, Lincoln P, Norden B: **Orientation and conformation of cell-penetrating peptide penetratin in phospholipid vesicle membranes determined by polarized-light spectroscopy**. *J Am Chem Soc* 2003, **125**:14214-14215.

32. Marrington R, Dafforn TR, Halsall DJ, Andrews T, Rodger A: **Micro volume Couette flow sample orientation for absorbance and fluorescence linear dichroism**. *Biophys J* 2004, in press.

Originally Published IVD Technology November/December 2004

Molecular Diagnostics

A linear dichroism technique for quantitative PCR applications

Research shows that an intrinsic signal from DNA has the potential to be an interference-free, specific, and label-free probe for real-time quantitative polymerase chain reaction.evelopment and progression.

David J. Halsall, Timothy R. Dafforn, Rachel Marrington, Eugene Halligan, and Alison Rodger

The high sensitivity and wide dynamic range of real-time polymerase chain reaction (RT-PCR) have made it the technique of choice for sensitive, specific nucleic acid quantitation, especially in the field of molecular diagnostics.[1] This article explores the novel application of flow-aligned linear dichroism (faLD) as a detection method for RT-PCR.

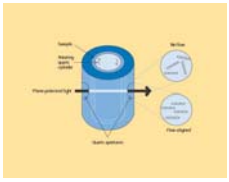


Figure 1. A schematic diagram of an LD Couette cell (click to enlarge).

Two intrinsic properties of the DNA helix are exploited by faLD: its ability to absorb ultraviolet (UV) light and its asymmetric structure. Consequently, unlike most current methods, faLD does not require costly exogenous labels or time-consuming post-PCR processing. Common interferences, such as those caused by artifactual PCR primer self-concatamerization (so-called primer-dimers), are effectively negated.[2] This is because, though they are optically active, they lack the asymmetry of the intended PCR product. Another advantage of the method is its ability to detect long PCR products. Also, it offers the possibility of detecting mutations concurrently with nucleic acid quantitation during the PCR process.

The Potential of faLD

Flow-aligned linear dichroism has potential as a method for detecting exclusively polymeric DNA.[3]

Molecules absorb light because the electric field of the radiation pushes their electrons in a particular direction at a particular wavelength. When all the molecules in a sample are oriented, the electrons in the sample all are characterized by the same preferred net displacement direction. Shining polarized light—that is, light whose electric field moves in only one direction—enables the absorbance to be varied.

Linear dichroism (LD) is the measure of the difference in absorbance between two perpendicular polarizations of light. This can be deconvoluted to yield information about the orientation of the absorbing units within the molecules. DNA bases exhibit intrinsic optical activity at 260 nm. These bases are arranged perpendicularly to the molecule's highly asymmetric double helix. As with logs in a river, when polymeric DNA is flowed through a narrow-walled channel, its helical axis lines up with the direction of flow. Because the electronic transitions of the DNA bases all involve movement of the electrons within the planes of the bases, light polarized perpendicularly to the helical axis is absorbed while light polarized along that axis is not. The absorbance of the parallel minus the absorbance of the perpendicularly polarized light gives a result that is the linear dichroism. The faLD of the DNA bases in polymeric duplex DNA thus is

negative.

The real attractiveness of using faLD for DNA detection is that it depends on the extent of orientation of the absorbing groups. With flow alignment, shear forces generated by fluid flow are used to align the asymmetric DNA. In aqueous buffer, the forces are too weak to induce any detectable degree of alignment until the duplex DNA is at least 200 base pairs (bp) in length. The flexibility of single-stranded DNA means that it, too, is essentially unaligned. Thus, when the degree of alignment is interrogated via the differential absorbance of two perpendicular, linearly polarized light beams, only duplex polymeric DNA will be seen.

This is particularly appealing, in prospect, for the detection of PCR products. The highly asymmetric products of the reaction have the potential to be aligned by shear force, whereas other optically active constituents or by-products of the PCR do not. The interferences that most commonly afflict RT-PCR systems would, therefore, effectively be rendered silent.

Apparatus and Technique

The rotating cuvette cell, called the Couette cell after its 19th-century inventor, is the apparatus most widely used to generate the shear forces required for flow-aligned LD spectroscopy (see Figure 1). The faLD spectrum of DNA is dominated by nucleotide absorbance at 260 nm, which is polarized perpendicularly to the axis of the helix and, as mentioned, has a negative value (see Figure 2). Molecules with absorbance moments more parallel to the alignment axis, such as DNA groove binders, generate a positive signal. Unaligned molecules have an LD of zero, which also is the case when the Couette cell is not rotating. A nonrotating cell thus provides a very simple baseline reading.

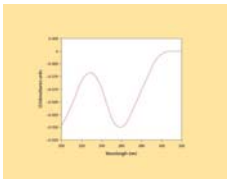


Figure 2. The faLD spectrum of calf thymus DNA (550 µmol) in aqueous buffer. Source: Biophysical Journal (2004). In press (click to enlarge).

The absolute magnitude of the LD signal depends not only on the angle between the absorbance axis of the optically active molecule and the axis of rotation, but also on the degree of alignment that can be achieved. If a molecule is too short or too flexible, or if the shear force generated by the Couette is weak, poor alignment results and the LD signal will be small. The shear force is a complex function of the design of the cell combined with the speed of rotation of the Couette, the size of the annular gap, and the viscosity of the medium. The LD signal obeys the Beer-Lambert law in that it is proportional to the concentration of alignable material within the sample (see Figure 3).

If the LD signal is expressed in ratio to the absorbance of the oriented sample, the result is a concentration-independent quantity that reflects the orientation of the absorbing species (in the case of DNA, bases nearly perpendicular to the helical axis)[4] and the length and rigidity of the polymer. This can be used to infer structural information about the nature of the LD-active species, as is discussed later in this article.

Recent developments in LD Couette design have reduced the cell volume such that new sample requirements and thermal-transfer characteristics make it appropriate to develop this technique as an RT-PCR system.[5] Sample volumes of 25 µl now are sufficient. For larger-volume applications, 600-µl and 3-ml cells are available. Routine PCR-reaction concentrations are ideal for LD measurements.

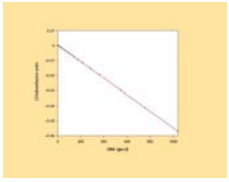


Figure 3. A graph displaying the dependence of faLD signal on DNA concentration, in this case the signal at 259 nmol of calf thymus DNA. Source: Biophysical Journal (2004). In press. [5] (click to

Couette cells have been in use since the 1890s. However, they were first developed into an instrument for flow dichroism in the 1960s.[6] Cell designs have used either a rotating inner and fixed outer cylinder or a rotating outer and fixed inner cylinder.6–9 The light path through the cell must be made of UV-transparent materials, to allow the optimal transmission of UV light that is necessary for LD of the DNA bases at 260 nm. Quartz is usually employed.

enlarge).

A microvolume Couette system has been designed specifically for low-volume applications such as the PCR performed with scarce material (see Figure 4).[5] The base plate of this system is designed to fit the sample compartment of a circular dichroism spectrophotometer made by Jasco Inc. (Easton, MD). Key system features that enable low-volume faLD measurements are a quartz capillary and a centrally mounted rod (see Figure 5). The capillary and rod are demountable for easy cleaning and are, in principle, disposable.

Adapted circular dichroism (CD) spectropolarimeters are usually employed for faLD experiments, as both linear and circular dichroism techniques require highly equivalent polarized light beams (see Figure 6). The photoelastic modulator of a CD spectropolarimeter needs to be adjusted for LD measurements in order to provide an oscillating half-wave plate with alternating horizontal and vertical polarizations. The wide light beam of most CD sources has to be focused onto the capillary housing (through lenses shown in Figure 4).

Detection of PCR Products

Because the LD signal is sensitive to both the concentration of double-stranded DNA in solution and its length, the increase in PCR product generated during a reaction can be readily detected using the negative LD signal at 260 nm. Currently available faLD apparatus is sufficiently sensitive to detect DNA generated by the PCR with no downstream processing (see Figure 7).

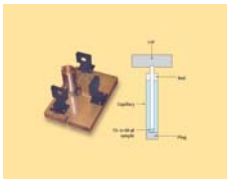


Figure 4. The design and construction of a microvolume faLD Couette system. (click to enlarge).

As has been predicted, a major advantage of faLD for the quantitation of PCR products is that primer-dimer artifacts, which plague conventional absorbance and fluorescence RT-PCR methods, are simply not an issue as they are too short to orient. PCR-generated primer-dimers in sufficient quantity to be detected by agarose gel electrophoresis with ethidium staining do not generate an faLD signal because they cannot be aligned by the shear forces generated by the apparatus.
Equipment currently under construction is designed to generate an LD signal during a PCR without the sample transfer that was required to generate the data represented in Figure 7.

A further advantage of faLD as a method for RT-PCR is the length of the PCR product, or amplimer, that can be effectively quantified. Most conventional RT-PCR systems recommend the use of amplimers less than 250 bp long. For applications such as detecting differentially spliced RNA transcripts, this can limit the specificity of the method. Figure 7 records the detection of a generated amplimer 1.3 kb long, but the technique can be used for amplimers even longer than this.

In fact, the signal-to-noise ratio is better with longer DNAs. This would be a disadvantage for applications that use amplimers shorter than 250 bp, which is often the case with high-sensitivity PCR, as short fragments can be amplified with greater efficiency. Preliminary data show that increasing the viscosity of the solvent by adding glycerol as a cosolvent allows shorter fragments to be aligned, but with the ultimate trade-off of signal generation by such artifacts as primer-dimers.

A formal appraisal of the analytical sensitivity of faLD has not yet been performed. This is because its PCR sensitivity is difficult to establish owing to the vast amplification that the PCR methodology involves. While the absolute sensitivity will be less than with a fluorescent label—that is, more amplimer is required to generate an LD signal than to generate a fluorescence signal—this is rarely an issue for most PCR applications because several more rounds of PCR will increase exponentially the amount of amplimer present. The benefits of being able to detect lower concentrations of amplimer with a fluorescence technique are offset by faLD's removal of the need for an exogenous label.

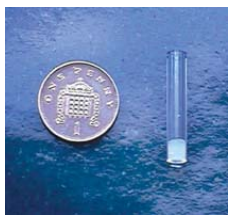


Figure 5. A microvolume faLD cell (click to enlarge).

Similarly, the analytical specificity of the faLD method is largely dependent on the choice of PCR primers. This is because the detection system usually

lends little contribution to specificity unless extra primers are introduced into the method. As mentioned previously, the faLD method greatly improves specificity in reactions where primer-dimer artifacts are an issue.

Practical Applications

The faLD method seems to have potential wide-scale clinical applicability. Apparent areas for use include viral titer determination, mutation detection, and multiplex PCR.

Viral Titer Determination. One practical application of faLD-PCR is to use the faLD signal to establish human immunodeficiency virus (HIV) titer in serum from patients infected with HIV (see Figure 8). The LD experiment involves simply measuring the signal of the post-PCR solution directly from the reaction without any purification. Because the LD is directly proportional to the amount of orientable DNA present, the LD signal gives a direct measure of the viral titer in the original sample.[5] The kinetic data provided by RT-PCR, as distinct from the end-time processing used in Figure 8, would provide additional quantitative information in such a situation. Thus, the DNA within the viral genome can be quantified without recourse to exogenous labels or downstream post-PCR processing.

Figure 6. The microvolume faLD Couette cell fitted into a Jasco J-715 CD spectropolarimeter (click to enlarge).

Mutation Detection. The distortion of the DNA helix that is caused by the introduction of a mismatched base pair has been exploited by a variety of techniques as a basis for mutation detection, the prototype being heteroduplex analysis.[10]

As discussed above, the LD signal is proportional to the extent of alignment of the DNA as summarized by the orientation parameter, which is itself dependent on the length and rigidity of the DNA. The introduction of even a single DNA mismatch has the potential to affect the net length of the DNA by, for example, kinking it, and also to affect its flexibility.

The distortion caused by a single base substitution in a 1.3-kb amplimer has been shown to affect the faLD spectra of a piece of DNA (see Figure 9).[11] Upon normalization of the faLD signals of wild-type and mutated DNA at 260 nm, the biggest change is observed at 230 nm, which is a minimum in the absorbance spectrum. In this case, the probe signal is, in fact, a scattering one, and is the result of the flow-aligned DNA changing shape due to the mutation.

The faLD technique therefore potentially combines amplimer quantitation and heterozygote detection in a homogeneous dye-free system. Its ability to detect a single base pair match in a 1.3-kb product far outperforms conventional methods, such as denaturing high-performance liquid chromatography and single-strand conformation polymorphism, that often are restricted in application to amplimers less than one third this size. The length of the PCR product may serve to amplify the distortion created by the base pair mismatch by virtue of the change in shape of the molecule.

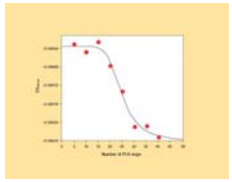


Figure 7. LD used to follow the progress of a PCR in which a 1.3-kb amplimer was generated using standard methods. After 20 cycles, the LD signal was significantly different from the initial value (click to enlarge).

FaLD offers the potential further advantage of being able to detect multiple mismatches in a single amplimer (See Figure 9). Its superiority over other mutation-screening methods lies in its lower cost and higher speed of analysis. However, the sensitivity of faLD for mutation detection remains to be proven.

Multiple Reactions. Fluorescence methods are well established, but they have reached their upper limit in terms of performing multiple quantitative PCR syntheses in a single reaction—so-called multiplex PCR. This is because fluorescence measurements require spectral windows for both exciting the dye and detecting the emitted photons. Owing to these constraints, most conventional systems are limited to maximum of five dyes.

The faLD-PCR could employ absorbance dyes, which have a far narrower spectral window. These dyes would become LD active only when incorporated into a PCR

amplimer, leading to an increase in both sensitivity and the ability to multiplex the LD-PCR technique. Although sacrificing the faLD advantage of not needing to optimize the PCR reaction with altered nucleotides, this could nevertheless be extremely attractive for particular assays.

More simply, differentiation between a more limited set of DNAs, for example, two or three viral DNAs of different lengths or degrees of flexibility or base content, could be undertaken by means of mathematical manipulation of the faLD spectrum and/or the real-time kinetics plots. This is because different DNAs exhibit different spectral envelopes, along with different polymerization kinetics and hydrodynamics (and, hence, orientation parameters). Such method development would have to proceed on a case-by-case basis, but it would be easy to undertake from a simple, automated real-time PCR run with wavelength scanning at each time point.

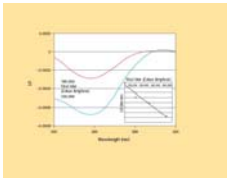


Figure 8. Flow-aligned linear dichroism used as a quantitative PCR method to determine viral titer, with the graph showing the faLD spectra of two PCRs performed on patients with different HIV viral titers as determined using the Cobas Amplicor method. The inset shows the dependence of the LD signal on titer. Primer sequences and reaction conditions were the same for both methods (click to enlarge).

DNA-Ligand Interactions

Not only can flow alignment of DNA generate an LD signal from the intrinsically optically active DNA bases, but it also is able to generate an LD signal from DNA-binding ligands that contain extrinsic chromophores. Binding captures the ligand in a fixed orientation relative to the DNA helix. Because free ligands will neither orient nor generate an LD signal, the LD signal that does result has the potential to interrogate binding kinetics. Furthermore, the faLD signal provides structural information pertaining to the DNA-ligand interaction. This is because it depends on the orientation of the structural elements within the ligand with respect to the DNA helical axis.

This has been the most widely used application of faLD to date. Flow-alignment LD is the technique of choice for discriminating between the two most common modes of ligand binding to DNA.[12] Intercalation between the DNA bases typically gives a negative faLD signal, as the absorbance dipoles of these ligands are typically oriented perpendicularly to the helical axis. In contrast, major or minor DNA groove binders can give positive LD signals. That is because these chromophores may have dipoles that can align parallel to the helical axis.

LD spectra of the intercalating ligand ethidium bromide, which has a negative signal, and the groove-binding ligand diaminophenyl indole, which shows a positive LD signal, may be compared (see Figure 10). The figure displays the typical 260-nm negative maxima from the DNA base pairs.

When the ligand involved is a protein, then complex, optically active structural elements within the polypeptide backbone, including helices, sheets, and optically active residues such as tryptophan, can all contribute to the LD signal. Detailed structural information about the DNA-RecA protein complex that is fundamental to the process of homologous recombination has been obtained using faLD.[13]

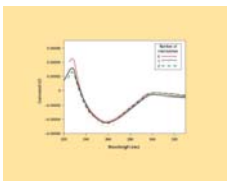


Figure 9. The detection of heteroduplex formation, with the graph showing the faLD spectra of PCR product with all combinations of two point mutations: none (the red dotted line), mutation A (black solid line), mutation B (orange dashed line), and both mutations (green dashed line).

Clearly, DNA-ligand interactions are of vital importance to biological processes such as DNA replication and transcription. Synthetic DNA ligands consequently have major clinical implications, particularly in the field of antitumor and antiviral drug development. The power of LD to investigate these interactions is reflected by an increasing number of publications detailing how LD was used to examine the binding modes of potential antitumor and antimicrobial compounds.[14–18] Since DNA-binding ligands, particularly proteins, can be notoriously difficult to obtain in large amounts, the development of the low-volume LD Couette system can be regarded as a major advance in widening the applicability of this method.

Conclusion

The maximum at 230 nm changes in intensity with respect to the number of mismatches as indicated (click to enlarge).

Flow-alignment linear dichroism is a promising detection system for quantitative PCR, both as an end-point method and for probing polymerization kinetics. It provides real-time PCR information because it involves neither exogenous labeling nor downstream processing of the PCR. This feature is unique among current methodologies.

The method has been shown to be robust with respect to monomer and primer-dimer interferences, and it can quantify longer amplimers better than many conventional systems. The faLD signal has the added potential to be used as a signal for heteroduplex analysis of sequence variants. The development of a low-volume Couette cell for quantitative PCR applications extends the applicability of this methodology, enabling it to be used to study DNA-ligand interactions when material is in short supply.

References

1. D Klein, “Quantification Using Real-Time PCR Technology: Applications and Limitations,” Trends in Molecular Medicine 8, no. 6 (2002): 257–260.

2. CT Wittwer et al., “Real-Time Multiplex PCR Assays,” Methods 25, no. 4 (2001): 430–442.

3. B Norden, M Kubista, and T Kurucsev, “Linear Dichroism Spectroscopy of Nucleic Acids,” Quarterly Review of Biophysics 25, no. 1 (1992): 51–170.

4. P Chou and WC Johnson, “Base Inclinations in Natural and Synthetic DNAs,” Journal of the American Chemical Society 115, no. 4 (1993): 1205–1214.

5. R Marrington et al., “Micro Volume Couette Flow Sample Orientation for Absorbance and Fluorescence Linear Dichroism,” Biophysical Journal (2004). In press.

6. A Wada and S Kozawa, “Instrument for the Studies of Differential Flow Dichroism of Polymer Solutions,” Journal of Polymer Science 2, part A (1964): 853–864.

7. P Oriel and J Schellman, “Studies of the Birefringence and Birefringence Dispersion of Polypeptides and Proteins,” Biopolymers 4 (1966): 469–494.

8. TR Dafforn et al., “Protein Fiber Linear Dichroism for Structure Determination and Kinetics in a Low-Volume, Low-Wavelength Couette Flow Cell,” Biophysics Journal 86, no. 1, part 1 (2004): 404–410.

9. C Lee and N Davidson, “Flow Dichroism of Deoxyribonucleic Acid Solutions,” Biopolymers 6 (1968): 531–550.

10. CF Taylor and GR Taylor, “Current and Emerging Techniques for Diagnostic Mutation Detection: An Overview of Methods for Mutation Detection,” Methods in Molecular Medicine 92 (2004): 9–44.

11. DJ Halsall, A Rodger, and TR Dafforn, “Linear Dichroism for the Detection of Single Base Pair Mutations,” Chemical Communications (Cambridge) 23 (2001): 2410–2411.

12. B Norden and T Kurucsev, “Analysing DNA Complexes by Circular and Linear Dichroism,” Journal of Molecular Recognition 7, no. 2 (1994): 141–155.

13. M Takahashi, K Morimatsu, and B Norden, “Structure of DNA-RecA Protein Complex, Intermediate of Homologous Recombination, Determined by Polarised-Light Spectroscopy,” Nucleic Acids Research Supplement no. 2 (2002): 9–10.

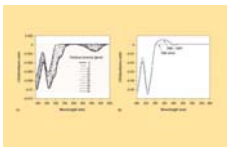


Figure 10. The faLD of DNA ligands with different binding modes: (a) 200-µmol calf thymus DNA with ethidium bromide, an intercalating ligand with negative LD signal, at 0–50 µmol in 10-mmol sodium cacodylate buffer (pH 7) and 10-mmol NaCl; and (b) 1000-µmol calf thymus DNA with 50-µmol diaminophenyl indole (DAPI), a groove-binding ligand with positive LD signal, in the same buffer. Source: Journal of the American Chemical Society (1993).[4] (click to enlarge).

14. S Hoet et al., “Alkaloids from Cassytha filiformis and Related Aporphines: Antitrypanosomal Activity, Cytotoxicity, and Interaction with DNA and Topoisomerases,” Planta Medica 70, no. 5 (2004): 407–413.

15. S Blanchard et al., “Synthesis of Mono- and Bisdihydrodipyridopyrazines and Assessment of Their DNA Binding and Cytotoxic Properties,” Journal of Medicinal Chemistry 47, no. 4 (2004): 978–987.

16. B Nguyen et al., “Characterization of a Novel DNA Minor-Groove Complex,” Biophysics Journal 86, no. 2 (2004): 1028–1041.

17. M Facompre et al., “Lamellarin D: A Novel Potent Inhibitor of Topoisomerase I,” Cancer Research 63, no. 21 (2003): 7392–7399.

18. O Novakova et al., “DNA Interactions of Monofunctional Organometallic Ruthenium(II) Antitumor Complexes in Cell-Free Media,” Biochemistry 42, no. 39 (2003): 11544–11554.

David J. Halsall, PhD, is a consultant clinical scientist at Addenbrooke’s Hospital (Cambridge, UK). He can be reached at djh44@hermes. cam.ac.uk. Timothy R. Dafforn, PhD, is an MRC Fellow in biosciences at the University of Birmingham (Birmingham, UK). He can be reached at t.r.dafforn@bham.ac.uk. Rachel Marrington, PhD, has just completed her doctoral studies at the University of Warwick (Warwick, UK). Eugene Halligan, PhD, is a clinical biochemist at Leicester Hospital NHS Trust (Leicester, UK). He can be reached at eh25@leicester.ac.uk. Alison Rodger, PhD, is the director of the MOAC (Molecular Organization and Assembly in Cells) doctoral training center and a reader in chemistry at the University of Warwick. She can be reached at a.rodger@warwick.ac.uk.

Design and DNA Binding of an Extended Triple-Stranded Metallo-supramolecular Cylinder

Carsten Uerpmann,[a] Jaroslav Malina,[b] Mirela Pascu,[b] Guy J. Clarkson,[b] Virtudes Moreno,[a] Alison Rodger,[b] Anna Grandas,*[a] and Michael J. Hannon*[b]

Abstract: A new tetracationic triple-stranded supramolecular cylinder is prepared from a bis(pyridylimine) ligand containing a diphenylmethane and two ketimine groups in the spacer. The cylinder is longer and slightly wider than the corresponding cylinder containing just diphenylmethane spacers. Inter-strand CH···π interactions are not observed and this affects the relay of the chiral information within the cylinder; a mixture of *rac* and *meso* isomers results, with the *meso* isomer being the dominant solution species and characterised in the solid state by crystallography. This new cylinder does bind to DNA as confirmed by induced circular dichroism signals in both the metal-to-ligand charge transfer (MLCT) and in-ligand bands of the cylinder. Flow linear dichroism demonstrates that the cylinder binds to DNA in a specific orientation(s) and is consistent with (major) groove-binding as seen for the shorter cylinder. Some DNA bending/coiling is observed but the effect is much less dramatic than observed for the cylinder with diphenylmethane spacers confirming that coiling is not solely a consequence of the tetracationic charge, but rather is related to the precise size and shape of the cylinder.

Keywords: bioinorganic chemistry · DNA recognition · helical structures · noncovalent interactions · supramolecular chemistry

Introduction

For most organisms, DNA encodes the molecular basis of life. Increasing amounts of information about the genetic sequence of ourselves and other species is becoming available, together with enhanced understanding of how this information is processed and regulated and how the processing is influenced by the environment. This offers many potential benefits for the quality of life. In particular, the ability to promote or to prevent the processing of specific genes will be crucial in the fight against many diseases, arising both from under- or over-processing of or damage to our own DNA and arising from external DNAs (as from viruses). For this reason drugs that act on DNA (at specific sequences) and can modulate DNA processing have great potential to form part of the armoury of drugs used within medicine in the 21st century.

Sequence-specific recognition of DNA by biomolecules occurs primarily through non-covalent interactions between the DNA and the surface motifs of proteins, and most commonly (though by no means exclusively) such recognition takes place in the DNA major groove.[1] This groove contains a more diverse pattern of hydrogen bond donor and acceptor units than the minor groove and its size and shape vary more with base sequence. DNA recognition by proteins is also sometimes associated with concomitant bending and coiling.[1] Synthetic agents that recognise DNA, in part because of their smaller size, most usually either intercalate between the base pairs[2] or bind in the minor groove.[3] We have recently demonstrated that metallo-supramolecular assembly may be used to prepare synthetic agents that are a similar size to the DNA recognition motifs found on proteins (such as zinc fingers or α-helices).[4] In particular we have developed metallo-supramolecular cylinders[4,5] that are a similar size and shape to zinc finger units and which do indeed seem to bind in the DNA major groove. Such

[a] Dr. C. Uerpmann, Prof. V. Moreno, Prof. A. Grandas
Departaments de Química Orgànica i Química Inorgànica, Universitat de Barcelona
Martí i Franquès 1–11, 08028 Barcelona (Spain)
Fax (+34) 933-397-878
E-mail: anna.grandas@ub.edu

[b] Dr. J. Malina, M. Pascu, Dr. G. J. Clarkson, Dr. A. Rodger, Dr. M. J. Hannon
Centre for Supramolecular and Macromolecular Chemistry
Department of Chemistry, University of Warwick
Gibbet Hill Road, Coventry, CV4 7 AL (UK)
Fax (+44) 2476-524-112
E-mail: m.j.hannon@warwick.ac.uk

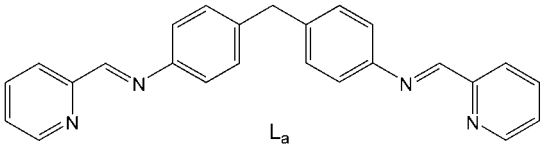
DOI: 10.1002/chem.200401054

metal-based assembly of these cylinders not only allows the nanoscale structure to be created but also imparts cationic charge (the cylinder is a tetracation). This is important since electrostatics will contribute significantly to the strength of the non-covalent binding of such species to the anionic DNA. Fascinatingly, the cylinders not only seem to target the major groove but also induce dramatic intramolecular DNA coiling which is unprecedented with synthetic agents and, in part, reminiscent of DNA coiling induced by histones in the cell nucleus. This is exciting since such coiling might represent a way to prevent gene processing. However, the precise manner in which the molecular-level major groove binding interaction of the cylinder leads to the observed macromolecular DNA coiling effect remains to be elucidated and represents a complex challenge because of the different size-scales on which these two events take place. To try to learn more about the essential design features, we have initiated a programme to investigate the effects of modifications to the basic cylinder structure on both the DNA binding and coiling. Herein we investigate a cylinder system in which we keep the same tetracationic charge as the original cylinder but increase the size of the cylinder unit.

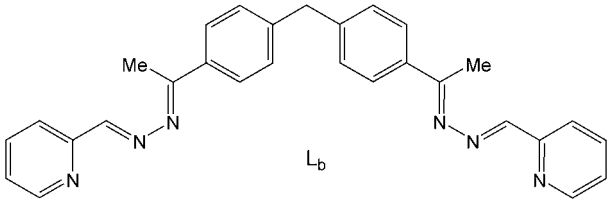
The interaction of metal complexes with DNA has attracted particular interest because cis-platin, and its analogues carboplatin, nedaplatin and oxaliplatin are the most widely used clinical anticancer agents and DNA is believed to be their target.[6] Cis-platin forms metal-nitrogen bonds to N7 of two adjacent purine bases (intra-strand GG and to a lesser extent GA lesions) which causes a kink (~45°) in the DNA which can then be recognised by nuclear HMG proteins. Mononuclear metal complexes that bind through non-covalent interactions (rather than forming bonds from the metal to the nitrogen atoms of the bases) have also been explored: For example, $[Cu(phen)_2]^+$ and "clip-phen" analogues bind in the minor groove and can cause oxidative backbone lesions.[7] $[Ru(phen)_3]^{2+}$ also binds non-covalently to DNA although the precise binding modes of its Δ and Λ enantiomers seems complex.[8] Functionalised analogues in which one phen ligand is extended to form an intercalation site have been developed and these metallo-intercalators may insert either from the major and minor grooves.[9] They have been applied to detect base pair mismatches[10] and rhodium(III) analogues have been used as photoactive footprinting agents.[11] Dinuclear analogues have been prepared which are proposed to thread through the DNA.[12] However none of these various non-covalent DNA-binding metal complexes are reported to exhibit the dramatic intramolecular DNA coiling observed with the cylinder. This may be because they are quite different in size and shape from the cylinders that we have developed, being smaller than the binuclear cylinders and affording smaller molecular surfaces which span only two-to-three DNA base pairs, and as such do not approach the size scale of nature's DNA recognition motifs.

Results and Discussion

To probe the effect of extending the cylinder structure, we made some modifications to the ligand structure. We were keen to try to retain structural similarity with the parent complex, $[Fe_2(L_a)_3]^{4+}$, while simultaneously increasing the dimensions. Our approach was to insert into the ligand design (L_b) ketimine spacer units between the pyridylimine



L_a



L_b

metal binding units at the ends of the ligand and the diphenylmethane central spacer unit. In this way we aimed to retain the basic structural features found in the original ligand L_a and its complex, but to extend the structure. While we have focused on pyridylimine binding sites in ligands for design of supramolecular architectures, Albrecht et al.[13] and Lehn et al.[14] have explored imine groups as *spacer units* in helicate design; the new design herein combines both approaches.

The ligand L_b was prepared by reaction of 4,4′-diacetyldiphenylmethane with two equivalents of hydrazine and subsequent condensation of the product with pyridine-2-carboxaldehyde to give the ligand. Reaction of the ligand L_b with iron(II) tetrafluoroborate in a chloroform–methanol mix at room temperature afforded the red complex $[Fe_2(L_b)_3][BF_4]_4$. The corresponding chloride salt could be prepared by an analogous route from iron(II) chloride. Electrospray mass spectrometry of an acetonitrile solution of the tetrafluoroborate salt revealed peaks corresponding to $\{Fe_2(L_b)_3(BF_4)_2\}^{2+}$, $\{Fe_2(L_b)_3F\}^{3+}$ and $\{Fe_2(L_b)_3\}^{4+}$ consistent with a dinuclear triple-stranded array.

The IR spectrum shows peaks corresponding to the coordinated ligand and the tetrafluoroborate counterion. The UV/Vis absorption spectrum revealed a band centered at 525 nm ($\varepsilon = 12\,000$) together with a shoulder at 485 nm ($\varepsilon = 10\,000$) corresponding to metal-to-ligand charge transfer (MLCT) transitions which are also observed in other iron(II) polypyridyl[15] and iron(II) pyridylimine complexes[4,5] and which confirm coordination of the iron(II) centre to three pyridylimine units in a low spin configuration. Acetonitrile solutions of the complex retain their red colour, arising from this visible MLCT transition, over a period of days indicating that the complex is stable in acetonitrile solution.

The ^{1}H NMR spectrum of the complex in acetonitrile solution reveals sharp peaks in the region 0–10 ppm confirming the diamagnetic nature of the complex. Two distinct sets of signals are observed in a ratio of about 7:1 indicating the presence of two solution species. Within a dinuclear triple-stranded formulation two configurations are possible: a helical or *rac* isomer in which the stereochemistry of the two metal centres is the same (and the isomer is thus chiral and is one of a pair of enantiomers) and a *meso* isomer in which the stereochemistry is different at the two metal centres rendering the structure achiral.[5,16–18] The central CH_2 resonance of the ligand can be used to distinguish these two isomers.[5,17] In the *rac* isomer the two protons of the CH_2 group are equivalent and give rise to a singlet, whereas in the *meso* isomer they are different and give rise to two doublets and we have previously observed this effect in the dinuclear double-stranded systems of ligand L_a.[17] For this triple-stranded dinuclear complex of L_b we similarly observe a pair of doublets and a singlet for the central CH_2 resonance and this enables us to identify the dominant species (85–90%) as the *meso* isomer and the minor component (10–15%) as the helical *rac* isomers. This contrasts with the iron(II) complex of L_a in which the sole products are the two enantiomeric helical *rac* isomers in both solution and the solid state. This change is presumably a consequence of the greater conformational and tortional flexibility introduced by the additional ketimine units in the spacer and is discussed further below.

Recrystallisation of the tetrafluoroborate salt from an acetonitrile solution by slow diffusion of benzene afforded red crystals which proved suitable for investigation by X-ray diffraction. The crystal structure (Figure 1) reveals a dinuclear

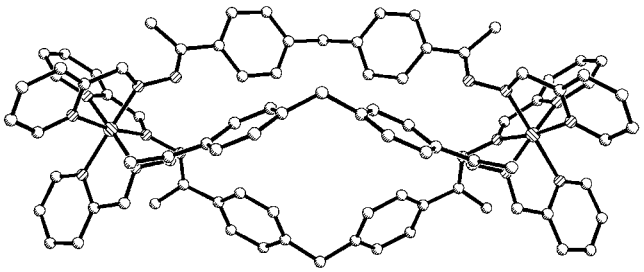


Figure 1. Crystal and molecular structure of the $[Fe_2(L_b)_3]^{4+}$ cation. Hydrogen atoms are omitted for clarity.

triple-stranded cylinder. However, in contrast to the iron(II) complex of the parent ligand L_a which is helical in both solution and the solid state, with ligand L_b the meso isomer crystallises in which the configuration of the two metal centres within the cylinder is opposite leading to an achiral (*meso*) structure. It is pertinent to contrast the structures of the two different cylinders obtained with L_a and L_b (Figure 2 and Figure 3).

As would be anticipated the longer spacer leads to an increase in the metal–metal separation (14.67 Å for the L_b cylinder, 11.52 Å for the L_a cylinder) and an increase of around

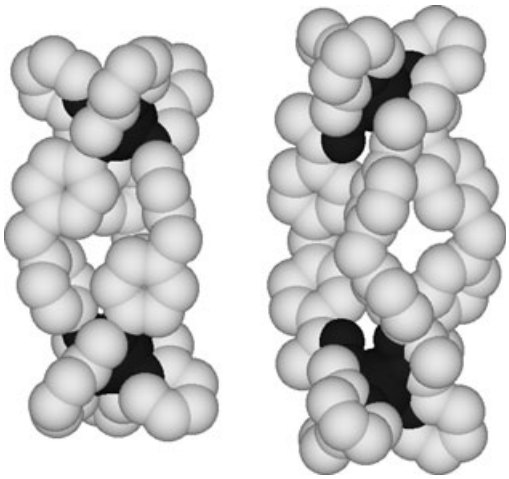


Figure 2. Space-filling representations of the cations *rac*-$[Fe_2(L_a)_3]^{4+}$ (left) and *meso*-$[Fe_2(L_b)_3]^{4+}$ (right). Hydrogen atoms are omitted for clarity.

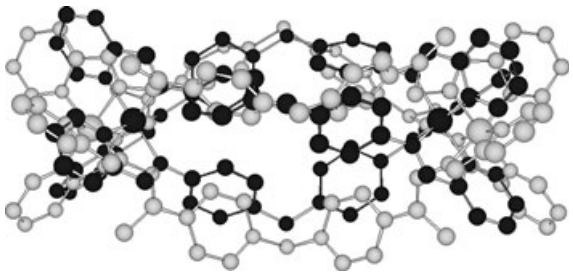


Figure 3. Overlay of the structures of the cations $[Fe_2(L_a)_3]^{4+}$ (dark) and $[Fe_2(L_b)_3]^{4+}$ (light). Hydrogen atoms are omitted for clarity.

12% in the overall length of the cylinder (C···C along the metal–metal vector: 20.38 Å for L_b, 17.34 Å for L_a: H···H 21.92 Å for L_b, 19.46 Å for L_a). The radius of the new L_b cylinder is also increased (measured at the central CH_2 unit 4.66 Å to C; 5.82 Å to H) and is approximately 8% larger than that of the corresponding L_a cylinder (radius 4.43 Å to C; 5.39 Å to H).

While in the complex of L_a the central aryl rings within the cylinder are all face–edge π-stacked together forming CH···π interactions, such interactions are absent in this longer cylinder (Figure 4). The pyridylimine units are planar (pyridylimine torsion angles 0.2°) and the phenyl and ketimine are also approximately coplanar (torsion angles ~23). Twisting within the ligand occurs primarily about the N–N bond between the imine and ketimine units (torsion angles ~85°).

The solid-state and solution structures are also pertinent to a recent report[19] by Albrecht et al. on the 'even-odd principle'[20] in helicate-box formation. In Albrecht's catechol-based systems he notes a tendency for systems with odd numbers of methylene units in the spacer to give *meso* isomers and those with even numbers to give helicates. He attributes this to the energetics associated with the conformation of the methylene chain. In a very recent study aimed at design of "two nanometer-dimensioned" helicates he has

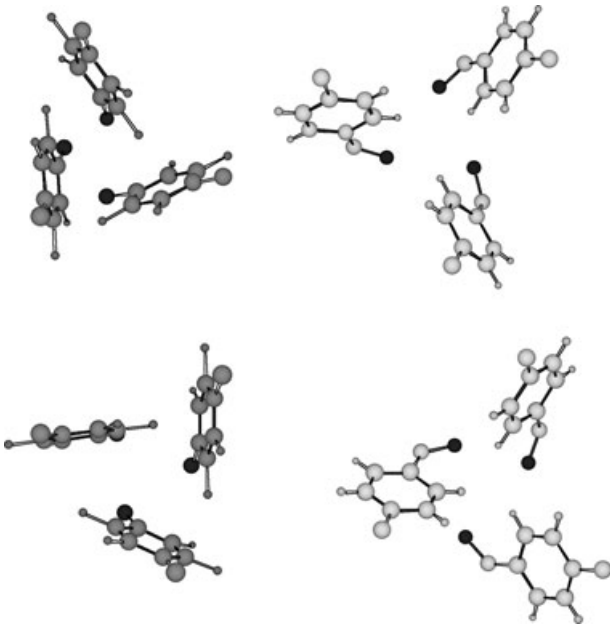


Figure 4. Comparison of the positions of the phenyl rings in the structures of the cations $[Fe_2(L_a)_3]^{4+}$ (left, dark shading) and $[Fe_2(L_b)_3]^{4+}$ (right, light shading) demonstrating the absence of the CH···π interactions in the L_b complex.

employed the diphenylmethane spacer and linked it through imine links to catechol binding units.[19] The result is a ligand in which the chelating units are positioned in a similar fashion to L_b and with titanium(IV) the dicatechol ligand gives a solid-state triple-stranded dinuclear *meso* isomer similar to the pyridylimine structure herein, and only slightly shorter (19.3 Å C···C along the metal-metal vector). To this extent the bis(pyridylimine) *meso* cylinder herein obeys Albrecht's "even-odd principle" as Albrecht's related bis-catechol. The previously reported L_a complex,[5] which is exclusively a triple-helix, is clearly at odds with the "even-odd" principle. Albrecht attributes this to the rigidity of that ligand system which enforces the same chirality at each metal centre.[19] We would concur, noting further that in the L_a system the extensive CH···π interactions at the centre of the complex will further increase the structural rigidity and thus facilitate the relay of the chiral information from one metal centre to the other. In L_b the additional ketimine units provide both additional conformational freedom and move the aryl rings of the spacer apart thereby removing that element of rigidity too. This presumably removes the constraints and allows the spacer conformational energetics associated with the "even-odd principle" to influence the stereochemistry. Nevertheless these constraints may not be completely relaxed as indicated by the presence of 10–15% of the helical *rac* isomer in solution.

DNA binding studies: The tetrafluoroborate salt $[Fe_2(L_b)_3][BF_4]_4$ is soluble in acetonitrile but not water. To investigate whether this salt could be used, the complex was dissolved first in a small amount of acetonitrile and then water was added to give an approximately 20% acetonitrile solution. Within a short time the compound precipitates to give a colourless solution. Consequently DNA binding studies were conducted with the chloride salt $[Fe_2(L_b)_3]Cl_4$. Solutions of the chloride salts, were prepared by dissolving in small amounts of ethanol and then adding water to give a 20% aqueous ethanol solution. We observed a slow change in solution colour from red to orange with time (accompanied by an increase in the MLCT absorption and some changes in the ligand bands). Consequently all solutions used in the spectroscopic investigations were freshly prepared and discarded after one hour.

Circular dichroism (CD) spectroscopy: Titrations of the complex into calf thymus DNA solution were carried out at constant concentrations of DNA (500 μM), NaCl (20 mM) and sodium cacodylate buffer (1 mM). The complex itself shows no CD signal. Any CD signals that arise in the spectroscopic regions of the complex are therefore a consequence of its interaction with the DNA. On addition of the complex to DNA, signals in the MLCT region of the complex (~480 and 520 nm) and in the region of the ligand bands (~315 nm) are observed confirming binding of the cylinder to DNA (Figure 5). These induced signals are simi-

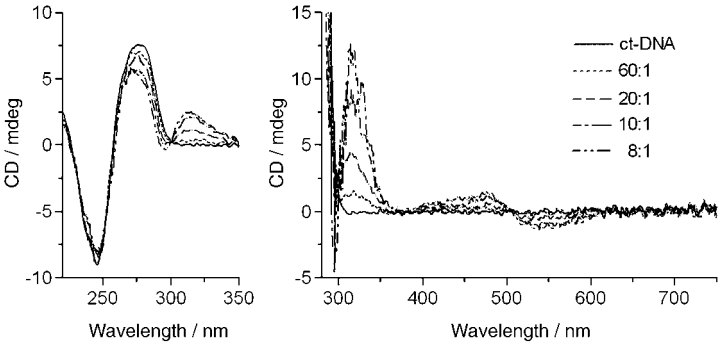


Figure 5. CD spectra of ct DNA (500 μM) in the presence of $[Fe_2(L_b)_3]Cl_4$. Mixing ratio ([DNA]: $[Fe_2(L_b)_3]Cl_4$) is indicated in the figure.

lar to those observed when the L_a triple-stranded cylinder binds to DNA[4] although in this case the sign of the band is inverted. The DNA CD bands (220–300 nm) confirm that the DNA remains in a B-DNA conformation. The increase in CD signal (314 nm) with concentration is linear up to five base pairs: one cylinder consistent with a single binding mode in this regime (from simple size considerations, the cylinder would span five base pairs if groove bound).

Flow linear dichroism (LD) spectroscopy: LD is the difference in absorption of light polarized parallel and perpendicular to an orientation direction. The technique is reliant on orientation of the sample and for DNA we achieve this in solution by orienting the biomacromolecule in a flow Couette cell with orientation through viscous drag.[21] The pri-

mary application of the technique has been to study drug-DNA interactions: if drugs are bound in a specific orientation(s) on the DNA then they too will become orientated, while drugs randomly bound to the DNA or free in solution will not be orientated. Since drugs which are not oriented will have identical absorptions of parallel and perpendicular polarised light, only those drugs which are oriented will exhibit an LD signal. Recently we have recognised that the technique can also be used to probe changes in the orientation of the biomacromolecule induced by drug binding events.[4] Thus for the L_a cylinders we observed dramatic losses of DNA orientation associated with the bending or coiling of the DNA by the cylinder. AFM imaging was able to confirm that this arose from an intramolecular DNA coiling effect.

To probe the binding of $[Fe_2(L_b)_3]Cl_4$, titrations were carried out at constant concentrations of DNA (500 μM), NaCl (20 mM) and sodium cacodylate buffer (1 mM). The presence of bands corresponding to cylinder transitions (~315, 520 nm) in the LD spectrum (Figure 6) confirms that the

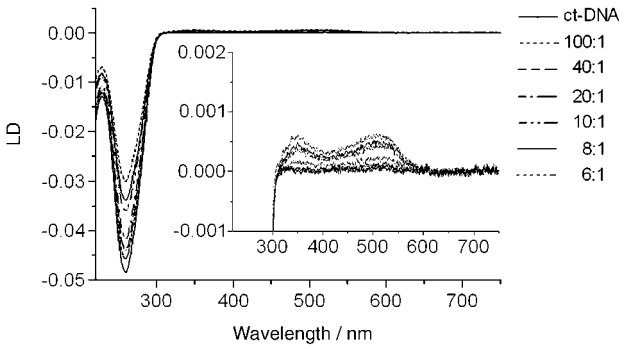


Figure 6. LD spectra of ct DNA (500 μM) in the presence of $[Fe_2(L_b)_3]Cl_4$. Mixing ratio ([DNA]: $[Fe_2(L_b)_3]Cl_4$) is indicated in the figure.

cylinder binds to DNA and does so in specific orientation(s). Some DNA bending or coiling is also observed as evidenced by the decrease in the DNA band at 260 nm. The loss of orientation is not as dramatic as that observed with the L_a iron(II) cylinder: In the LD of the DNA region the parent *P* enantiomer of the L_a cylinder causes a 75% loss of orientation at a 15:1 ratio and the *M* enantiomer is even more aggressive in its coiling causing a 95% loss at a 20:1 ratio. This new L_b cylinder shows a loss of only about 40% at 6:1.

Thermal stability: Thermal stability experiments (monitored by measuring cylinder absorbance at 476 nm) indicated that the complex was degraded as the temperature was increased above ambient and that the presence of DNA had no significant stabilising effect on the complex. For this reason, it is unsurprising that the cylinder also exhibited no effect on the melting temperature of ct-DNA.

Conclusion

The additional ketimine groups introduced into the spacer unit of L_b increase the separation of the two metal binding sites in the ligand and, as expected, extend the length of the cylinder created. However they also move the phenylene groups on different ligand strands apart and these phenylenes no longer form the inter-strand CH···π interactions observed in the complex of L_a. This affects the complex in several ways. The removal of the inter-strand interactions makes the relay of the chiral information from one metal centre to the other less effective and a mixture of *rac* and *meso* isomers results, with the *meso* isomer being dominant. We have recently demonstrated that inter-strand interactions are important in obtaining exclusively *rac* isomers in the related *double*-stranded systems.[17a] The diameter of the cylinder is also increased with respect to the L_a cylinder. The stability of the complex in aqueous solution is reduced, supporting the view that the inter-strand CH···π interactions do contribute to the stability of the complexes of ligand L_a.

This new cylinder does bind to DNA as evidenced by the observation of induced CD signals in the both the MLCT and in-ligand bands of the cylinder. The induced signals are similar to those observed for the L_a complex although opposite in sign. We speculate that this may be related to the fact that the L_a cylinder is *rac*, whereas the L_b cylinder is predominantly the *meso* isomer and contains metal centres with opposite configurations which will have diferent transition orientations. The lower aqueous stability of the L_b cylinder makes it difficult to explore this further. The flow LD experiment confirms that the L_b cylinder binds to DNA in a specific orientation(s) (as the L_a cylinder). The sign of the cylinder MLCT LD signal precludes this being an intercalative binding mode (consistent with the cylinder's large size and shape) but is consistent with a (major) groove-binding mode as seen for the L_a cylinder.

Some DNA bending/coiling is observed but the effect is much less dramatic than observed for the L_a cylinder. It is clear from this result that the DNA coiling observed with the L_a cylinder is not solely a consequence of the tetracationic charge, but rather is related to the precise size and shape of the cylinder. We are currently investigating other related cylinder designs to try to probe further the essential features of the design and their impact on the DNA recognition and coiling.

Experimental Section

General: All reagents and solvents were purchased from commercial sources (Aldrich, Fluorochem, Acros, Fluka) and used without further purification. NMR spectra were recorded on Varian Gemini 2000 (200 MHz) or Brüker DPX 400 (400 MHz) instruments at 298 K using standard Varian or Brüker software. FAB mass spectra were recorded on a Micromass VG-Quattro System or a Micromass AutoSpec spectrometer using 3-nitrobenzyl alcohol as matrix. ESI mass spectra of the complex were recorded on a Micromass Quatro II (low-resolution triple quadrupole mass spectrometer) instrument at the EPSRC National Mass Spec-

trometry Centre, University of Wales, Swansea and of the ligand on an Micromass ZQ System at U. Barcelona. Infrared spectra (solid pellets or films prepared by evaporation of solutions mounted on NaCl plates) were measured with a Perkin Elmer Paragon 1000 FTIR spectrometer or a Nicolet 510 FT-IR Spectrometer. UV/Vis measurements were made using a Jasco V-550 spectrophotometer.

Ligand L_b: 4,4′-Diacetyldiphenylmethane (0.508 g, 2 mmol) was dissolved in methanol (15 mL). Hydrazine monohydrate (0.210 g, 4.2 mmol) was added slowly over a period of 15 minutes and the mixture was stirred at room temperature overnight. The white precipitate of the bishydrazone formed (0.450 g, 1.6 mmol, yield 80%) was collected by filtration, washed with small portions of cold methanol (2×5 mL) and dried in vacuo.

^{1}H NMR (200 MHz, $CDCl_3$): $\delta = 7.56$ (d, $J = 8.2$ Hz, 4H; Ph), 7.16 (d, $J = 8.2$ Hz, 4H; Ph), 5.32 (s, 4H; NH_2), 3.98 (s, 2H; CH_2), 2.11 ppm (s, 6H; CH_3); ^{13}C NMR (50 MHz, $CDCl_3$): δ 11.7, 41.3, 125.5, 128.8, 140.8, 147.3 ppm; MS (+ve FAB, NBA): *m/z*: 281 $[M+H]^+$; IR (film): $\tilde{\nu} =$ 3354 s, 3220 s, 3083 m, 1654 m, 1604 m, 1509 m, 1436 m, 1368 m, 1335 m, 1108 m cm^{-1}.

The bis(hydrazone) (0.135 g, 0.48 mmol) and pyridine-2-carboxaldehyde (0.107 g, 1 mmol) were mixed in methanol (15 mL). After adding several drops of acetic acid the clear solution was stirred at room temperature overnight. The resulting yellow precipitate was collected by filtration and dried in vacuo (0.202 g, 0.44 mmol, yield 91%).

$R_f = 0.31$ (silica; CH_2Cl_2:MeOH 98:2); ^{1}H NMR (200 MHz, $CDCl_3$): $\delta =$ 8.68 (d, $J = 4.4$ Hz, 2H; py-H6), 8.41 (s, 2H; N=CH), 8.14 (d, $J = 7.8$, , 2H; py-H3), 7.86 (d, $J = 8.4$ Hz, 4H; Ph), 7.73 (td, $J = 7.5$, 1.8 Hz, 2H; py-H4), 7.35 (ddd, $J = 7.2$, 4.8, 1.1 Hz, 2H; py-H5), 7.25 (d, $J = 8.2$ Hz, 4H; Ph), 4.08 (s, 2H; CH_2), 2.47 ppm (s, 6H; CH_3); ^{13}C NMR (50 MHz, $CDCl_3$): δ: 15.4, 41.6, 121.5, 124.5, 127.2, 129.0, 136.4, 143.0, 149.6, 153.6, 157.1, 163.9 ppm; IR (film): $\tilde{\nu} =$ 3095 m, 1649 m, 1610 s, 1436 m, 1369 m, 1343 m, 1115 m cm^{-1}; MS (+ve FAB, NBA): *m/z*: 459 $[M+H]^+$; MS (+ve ESI): *m/z*: 459 $[M+H]^+$; elemental analysis calcd(%) for $C_{29}H_{26}N_6$: C 76.0, H 5.7, N 18.3; found: C 76.0, H 5.7, N 18.1.

Iron(II) complex: L_b (0.023 g, 0.05 mmol) and iron(II) tetrafluoroborate (0.025 g, 0.075 mmol) in a mixture chloroform–methanol (1:1) were stirred for 6 h at room temperature. The red precipitate was collected by vacuum filtration, washed with methanol and dried in vacuo under P_4O_{10} (0.021 g, 70%). X-ray quality, red crystals were obtained by slow diffusion of benzene into a solution of the complex in acetonitrile. The chloride salt was prepared by an analogous route from iron(II) chloride.

^{1}H NMR (400 MHz, CD_3CN, 298 K): $\delta = 8.31$ (s, 2H; H^{imine} *helix*), 8.25 (s, 14H; H^{imine} *box*), 8.18 (m, 16H; H^4), 8.09 (m, 16H; H^3), 7.79 (d, $J =$ 8.3 Hz, 28H; Ph *box*), 7.66 (d, $J = 8.0$ Hz, 4H; Ph *helix*), 7.59 (m, 32H; H^5& H^6), 7.33 (d, $J = 8.3$ Hz, 28H; Ph *box*), 7.20 (d, $J = 8.0$ Hz, 4H; Ph *helix*), 4.09 (s, 2H; CH_2 *helix*), 4.06 (d, $J = 23$ Hz, 7H; CH_2 *box*), 3.96 (d, $J = 23$ Hz, 7H; CH_2 *box*), 1.27 (s, 6H; CH_3 *helix*), 1.20 (s, 42H; CH_3 *box*); ^{1}H NMR (400 MHz, CD_3OD, 298 K, chloride salt): $\delta = 8.58$ (s, 2H; H^{imine} *helix*), 8.532 (s, 14H; H^{imine} *box*), 8.30 (m, 16H; H^4), 8.24 (m, 16H; H^3), 7.98 (d, $J = 8.0$ Hz, 28H; Ph *box*), 7.84 (d, $J = 7.8$ Hz, 4H; Ph *helix*), 7.75 (m, 32H; H^5& H^6), 7.37 (d, $J = 8.0$ Hz, 28H; Ph *box*), 7.20 (d, $J =$ 8.0 Hz, 4H; Ph *helix*), 4.13 (s, 2H; CH_2 *helix*), 3.97 (d, $J = 14$ Hz, 7H; CH_2 *box*), 3.88 (d, $J = 14$ Hz, 7H; CH_2 *box*), 1.34 (s, 48H; CH_3 *helix* & CH_3 *box*); IR (solid): $\tilde{\nu} = 1600$ m, 1562w,1472 m, 1439w, 1413w, 1371w, 1306 m, 1187w, 1162w, 1053 s, 919w, 813w, 777 s, 760 s, 686w cm^{-1}; UV/Vis (MeCN): $\lambda = 525$ ($\varepsilon = 12000$), 485 nm ($\varepsilon = 10000$), 296 nm ($\varepsilon =$ 138000), 260 nm ($\varepsilon = 95000$); MS (+ve ESI, MeCN): *m/z*: 372 $\{Fe_2(L_2)_3\}^{4+}$ (100%), 502 $\{Fe_2(L_2)_3F\}^{3+}$ (10%), 830 $\{Fe_2(L_2)_3(BF_4)_2\}^{2+}$ (5%).

X-ray crystallography: Crystal structure data for $C_{43.8}H_{42.8}B_{1.33}F_{5.33}Fe_{0.67}N_{7.4}O_{0.5}$, $M_r = 833.83$, hexagonal, space group *P*6(3)/*m*, $a = 13.791(3)$, $b =$ 13.791(3), $c = 44.622(10)$ Å, $U = 7349(3)$ Å^3, $Z = 6$, $\rho_{calcd} = 1.130$ g cm^{-3}, $\mu(Mo_{K\alpha}) = 0.270$ mm^{-1}. Character: purple blocks $0.28 \times 0.28 \times 0.04$ mm. A total of 30463 reflections were measured, 3620 unique[R(int) = 0.1462]. Goodness-of-fit on F^2 was 1.051. $R1$ [for 2373 reflections with $I > 2\sigma(I)$] = 0.0883, $wR2 = 0.2225$. Data/restraints/parameters 3620/22/327. Crystal data was collected at 180(2) K with a Siemens-SMART-CCD diffractometer[22] equipped with an Oxford Cryosystem Cryostream Cooler.[23] Refinements used SHELXTL.[24] Systematic absences indicated space group *P*6(3)/*m* or *P*6(3). The former was chosen on the basis of intensity statistics and shown to be correct by successful refinement. The structure was solved by direct methods with additional light atoms found by Fourier methods. Hydrogen atoms were added at calculated positions and refined by using a riding model with freely rotating methyl groups. Anisotropic displacement parameters were used for all non hydrogem atoms; hydrogen atoms were given isotropic displacement parameters equal to 1.2 (or 1.5 for methyl hydrogen atoms) times the equivalent isotropic displacement parameter of the atom to which the hydrogen atom is attached.

The asymmetric unit contains half a ligand, an iron, a benzene molecule, two partially occupied acetonitriles (N200–C202, 50% occupancy and N400–C402, 20% occupancy) two partially occupied water molecules and two very disordered BF_4 counter ions (as detailed below). Fe1 sits on special position 4f on a three-fold axis. C18, the central methylene of the ligand lies on a mirror plane, special position 6h. B10 sits on special position 4f on a three-fold axis and has the fluorine atoms disordered over two positions in a 3:1 ratio. The major position F11/F12 has F11 on the same three-fold axis and F12 in a general position. The minor component F11 A/F12 A has F11 A on the three-fold axis and F12 A in a general position. B20 sits on special position 4e on a three fold axis and has very disordered fluorine atoms over at least three positions. This was modelled by placing four fluorine atoms (F21–F24) on B20 and making the sum of their occupancies add up to a BF_4. Acetonitrile N400–C402, as well as being only partially occupied lies on the mirror plane, special position 6h. Some additional electron density was modelled as partially occupied water molecules O600 and O700. O600 lies on special position 4f on a three-fold axis and is 50% occupied. O700 lies on special position 4e on a three-fold axis and is 25% occupied. The partial solvent occupancies and no hydrogen atoms on the water molecules explains the fractional formula.

CCDC-252893 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Biophysical experiments: *Circular dichroism*: Spectra were collected in 1 cm pathlength cuvettes (280–850 nm) or 2 mm pathlength cuvettes (220–350 nm) using a Jasco J-715 spectropolarimeter. Spectroscopic titrations were performed in which CD and UV/Vis absorbance spectra were collected. Titrations were carried out at constant concentrations of DNA (500 μM), NaCl (20 mM) and sodium cacodylate buffer (1 mM). The DNA:metal complex ratio was decreased during the titration series by incrementing the concentration of metal complex in the cuvette from 0–62.5 μM. Titrations were performed so as the concentrations of DNA, NaCl and sodium cacodylate buffer in the cuvette remained unaltered.[25]

Flow linear dichroism: Flow *LD* spectra were collected by using a flow couette cell in a Jasco J-715 spectropolarimeter adapted for LD measurements. Long molecules, such as DNA (minimum length of ~250 base pairs), can be orientated in a flow couette cell. The flow cell consists of a fixed outer cylinder and a rotating solid quartz inner cylinder separated by a 0.5 mm wide gap giving a total 1 mm pathlength.[21]

Thermal stability measurements: Thermal stability experiments used a thermostatic cell-changer UV/Vis spectrophotometer on a Cary1E spectrophotometer in masked 1.2 mL cuvettes, monitoring cylinder absorbance at 476 nm.

Acknowledgements

We thank the European Union (MARCY RTN, HPRN-CT-2002–00175; Marie Curie fellowships: M.P. HPMT-GH-01–00365–11; J. M. HPMF-CT-2002–01856); the Spanish MEC (BQU2001–3693-C02–01) and the Generalite de Catalunya (2001SGR49 and Centre de Referencia de Biotechnologia) for funding. This work was conducted in the context of COST D20 WG D20/0010/02 on non-covalent DNA recognition. We thank EPSRC and Siemens Analytical Instruments for grants in support of the diffractometer and the EPSRC National Mass Spectrometry Service Centre,

Swansea for recording electrospray mass spectra. M.J.H. is the Royal Society of Chemistry Sir Edward Frankland Fellow 2004–2005.

[1] See for example: J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th ed., Freeman, New York, **2002**; C. Branden, J. Tooze, *Introduction to Protein Structure*, 2nd ed., Garland, New York, **1999**; S. Neidle *Nucleic Acid Structure and Function*, OUP, Oxford, **2002**; R. E. Dickerson, *Nucleic Acids Res.* **1998**, *26*, 1906–1926.

[2] See for example: B. C. Baguley, *Anti-Cancer Drug Des.* **1991**, *6*, 1–35; S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry* University Science Books, Mill Valley, CA, **1994**; P. B. Glover, P. R. Ashton, L. J. Childs, A. Rodger, M. Kercher, R. M. Williams, L. De Cola, Z. Pikramenou, *J. Am. Chem. Soc.* **2003**, *125*, 9918–9919.

[3] P. B. Dervan, B. S. Edelson, *Curr. Biol. Curr. Opin. Struc. Biol.* **2003**, *13*, 284–299; P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.

[4] a) M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 880–884; b) I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, M. Rodger, J. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074; c) E. Moldrheim, M. J. Hannon, I. Meistermann, A. Rodger, E. Sletten, *J. Biol. Inorg. Chem.* **2002**, *7*, 770–780; d) M. J. Hannon, A. Rodger, *Pharmaceutical Visions* **2002**, autumn issue, 14.

[5] M. J. Hannon, C. L. Painting, J. Hamblin, A. Jackson, W. Errington, *Chem. Commun.* **1997**, 1807–1808.

[6] *Cisplatin, Chemistry and Biochemistry of a Leading Anti-Cancer Drug* (Ed.: B. Lippert) Wiley-VCH, Weinheim, Germany, **1999**; J. Reedijk, *Chem. Commun.* **1996**, 801–806; Z. J. Guo, P. J. Sadler, *Adv. Inorg. Chem.* **2000**, *49*, 183–306 and references therein.

[7] D. S. Sigman, T. W. Bruice, A. Mazumder, C. L. Sutton, *Acc. Chem. Res.* **1993**, *26*, 98–104; M. Pitié, B. Donnadieu, B. Meunier, *Inorg. Chem.* **1998**, *37*, 3486–3489.

[8] C. V. Kumar, J. K. Barton, N. J. Turro, *J. Am. Chem. Soc.* **1985**, *107*, 5518–5523; C. Hiort, B. Nordén, A. Rodger, *J. Am. Chem. Soc.* **1990**, *112*, 1971–1982; D. Z. Coggan, I. S. Haworth, P. J. Bates, A. Robinson, A. Rodger, *Inorg. Chem.* **1999**, *38*, 4486–4497.

[9] K. E. Erkkila, D. T. Odom, J. K. Barton, *Chem. Rev.* **1999**, *99*, 2777–2795; I. Greguric, J. R. Aldrich-Wright, J. G. Collins, *J. Am. Chem. Soc.* **1997**, *119*, 3621–3622; E. Tuite, P. Lincoln, B. Norden, *J. Am. Chem. Soc.* **1997**, *119*, 239–240; P. Lincoln, E. Tuite, B. Norden, *J. Am. Chem. Soc.* **1997**, *119*, 1454–1455.

[10] K. E. Erkkila, B. P. Hudson, J. K. Barton, D. C. Rees, *Nat. Struct. Biol.* **2000**, *7*, 117–121; U. Schatzschneider, J. K. Barton, *J. Am. Chem. Soc.* **2004**, *126*, 8630–8631; E. Ruba, J. R. Hart, J. K. Barton, *Inorg. Chem.* **2004**, *43*, 4570–4578.

[11] H. Junicke, J. R. Hart, J. Kisko, O. Glebov, I. R. Kirsch, J. K. Barton, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3737–3742; B. A. Jackson, J. K. Barton, *J. Am. Chem. Soc.* **1997**, *119*, 12986–12987.

[12] B. Önfelt, P. Lincoln, B. Nordén, *J. Am. Chem. Soc.* **2001**, *123*, 3630–3637.

[13] See for example M. Albrecht, I. Janser, S. Kamptmann, P. Weis, B. Wibbeling, R. Fröhlich, *Dalton Trans.* **2004**, 37–43.

[14] K. M. Gardinier, R. G. Khoury, J. M. Lehn, *Chem. Eur. J.* **2000**, *6*, 4124–4131; R. Stiller, J. M. Lehn, *Eur. J. Inorg. Chem.* **1998**, 977–982.

[15] E. C. Constable, *Adv. Inorg. Chem. Radiochem.* **1986**, *28*, 69–121; E. C. Constable, *Adv. Inorg. Chem.* **1989**, *34*, 1–63.

[16] L. J. Childs, M. J. Hannon, *Supramol. Chem.* **2004**, *16*, 7–22; M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3497; C. Piguet, G. Bernardinelli, G. Hopfgartner, *Chem. Rev.* **1997**, *97*, 2005–2062.

[17] a) L. J. Childs, M. Pascu, A. J. Clarke, N. W. Alcock, M. J. Hannon, *Chem. Eur. J.* **2004**, *10*, 4291–4300; b) M. J. Hannon, C. L. Painting, N. W. Alcock, *Chem. Commun.* **1999**, 2023–2024; c) L. J. Childs, M. J. Hannon, N. W. Alcock, *Angew. Chem.* **2002**, *114*, 4418–4420; *Angew. Chem. Int. Ed.* **2002**, *41*, 4244–4247.

[18] A. Bilyk, M. M. Harding, P. Turner, T. W. Hambley, *J. Chem. Soc. Dalton Trans.* **1994**, 2783–2790; C. O. Dietrich-Buchecker, J. F. Nierengarten, J. P. Sauvage, N. Armaroli, V. Balzani, L. DeCola, *J. Am. Chem. Soc.* **1993**, *115*, 11237–11244.

[19] M. Albrecht, I. Janser, H. Houjou, R. Fröhlich, *Chem. Eur. J.* **2004**, *10*, 2839–2850.

[20] M. Albrecht, *Chem. Eur. J.* **2000**, *6*, 3485–3489.

[21] A. Rodger, B. Norden, *Circular and Linear Dichroism*, Oxford University Press, Oxford, **1997**.

[22] SMART user's manual, Siemens Industrial Automation Inc., Madison, WI, **1994**.

[23] J. Cosier, A. M. Glazer, *J. Appl. Crystallogr.* **1986**, *19*, 105–107.

[24] G. M. Sheldrick, *Acta Crystallogr. Sect. A* **1990**, *46*, 467–473.

[25] A. Rodger, *Methods Enzymol.* **1993**, *226*, 232–258.

Received: October 15, 2004
Published online: January 25, 2005

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

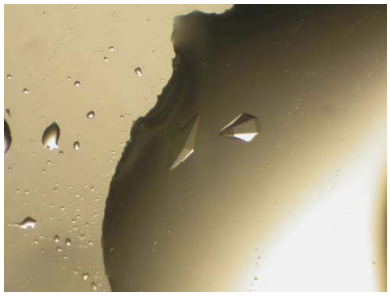
Expression, purification and crystallization of the cell-division protein YgfE from *Escherichia coli*

Stephen G. Addinall,[a] Kenneth A. Johnson,[b] Timothy Dafforn,[c] Corinne Smith,[d] Alison Rodger,[e,f] Raul Paco Gomez,[d,f] Katherine Sloan,[a] Anne Blewett,[d] David J. Scott[g] and David I. Roper[d]*

[a]Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, England, [b]Department of Biochemistry and Biophysics, The Arrhenius Laboratories for Natural, Sciences, Stockholm University, SE-106 91 Stockholm, Sweden, [c]Department of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, England, [d]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, England, [e]Department of Chemistry, University of Warwick, Coventry CV4 7AL, England, [f]MOAC Doctoral Training Centre, University of Warwick, Coventry CV4 7AL, England, and [g]National Centre for Macromolecular Hydrodynamics, School of Biosciences, University of Nottingham, Sutton Bonington, Leicestershire LE12 5RD, England

Correspondence e-mail: david.roper@warwick.ac.uk

Received 21 December 2004
Accepted 3 February 2005
Online 24 February 2005



An open reading frame designated b2910 (ygfE) in the *Escherichia coli* K12-MG1655 genome sequence, identified as a possible homologue to the cell-division protein ZapA, was cloned into the high-expression plasmid pETDuet-1 and overexpressed in *E. coli* BL21 (DE3)-AI. The protein was purified in three steps to 99% purity. Crystals were obtained by the hanging-drop vapour-diffusion method at 291 K from a wide range of screened conditions, but principally from solutions containing 0.1 *M* HEPES pH 7.0, 18% PEG 6000, 5 m*M* $CaCl_2$. Diffraction data to 1.8 Å were collected at the European Synchrotron Radiation Facility (ESRF). The crystals belong to space group $P6_122$ or $P6_522$, with unit-cell parameters $a = 53.8$, $b = 53.8$, $c = 329.7$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$.

1. Introduction

The ancestral tubulin protein FtsZ has a crucial role in the process of bacterial cell division. FtsZ polymerizes into a ring (the 'Z-ring') around the middle of pre-divisional cells on the inner surface of the cytoplasmic membrane. The Z-ring then constricts, driving cell division (Addinall & Holland, 2002; Lutkenhaus & Addinall, 1997; Margolin, 2000). FtsZ is extensively conserved throughout bacterial species and plants (where it is required for plastid division). FtsZ has also been implicated in the division of some archaeal species and the mitochondria of some eukaryotes (Vaughan *et al.*, 2004). As a cytoskeletal element with prokaryotic origins the Z-ring represents a fascinating topic of study and as a major player in a fundamental process for bacterial life FtsZ represents an exciting potential target for novel antibacterial compounds.

The mechanisms by which Z-rings form and constrict are poorly characterized and the detailed *in vivo* structure of the Z-ring is unknown. *In vitro* studies reveal that FtsZ is a GTPase which can polymerize into linear polymers in a GTP-dependent manner (Romberg & Levin, 2003), and that these polymers form lateral associations to a greater or lesser degree depending on polymerization conditions. Measurements indicate that cells contain sufficient FtsZ for the Z-ring to be composed of multiple protofilaments (Lu *et al.*, 1998; Feucht *et al.*, 2001). This is likely to be the case since cells containing GFP-FtsZ fusion protein incorporate approximately 30% of the GFP fluorescence into the Z-ring (Anderson *et al.*, 2004). Therefore, lateral interactions between FtsZ polymers are likely to be important in Z-ring function.

Two FtsZ accessory proteins have been identified that promote lateral associations between FtsZ polymers. ZipA (from *Escherichia coli*) is a membrane-spanning protein with a cytoplasmic domain that interacts with the FtsZ C-terminus (Hale & de Boer, 1997). This domain of ZipA induces bundling of FtsZ polymers *in vitro* (Hale *et al.*, 2000; RayChaudhuri, 1999) and its structure has been determined in complex with a small FtsZ peptide (Moy *et al.*, 2000). ZapA from both *Bacillus subtilis* and *Pseudomonas aeruginosa* has been demonstrated to bind and to induce bundling between FtsZ polymers (Gueiros-Filho & Losick, 2002; Low *et al.*, 2004), and the crystal structure of tetrameric *P. aeruginosa* ZapA has been solved (Low *et al.*, 2004). We recently demonstrated that the *E. coli* YgfE protein binds to and bundles FtsZ polymers, inhibits the FtsZ GTPase and adopts a dimer–tetramer equilibrium in solution (Small *et al.*,

submitted work). YgfE is thus a functional homologue of ZapA as suggested by previous sequence and immunofluoresence data (Gueiros-Filho & Losick, 2002).

2. Methods and materials

2.1. Cloning, overexpression and purification

The gene encoding *E. coli* YgfE (b2910; 109 amino acids, 12 594 Da) was amplified by PCR from *E. coli* K12-MG1655 chromosomal DNA with primers incorporating an *Eco*RI site at the 5′ end of the gene and a *Hin*dIII site at the 3′ end. The PCR fragment was digested with these enzymes and ligated into a pETDuet-1 vector (Novagen), also digested with *Eco*RI and *Hin*dIII. This strategy provides YgfE with an amino-terminal six-histidine tag. After selection of recombinants containing the cloned gene by restriction digest, *E. coli* BL21 (DE3) AI (Invitrogen) was transformed with the resulting recombinant plasmid pETDuet-*ygfE*. Cells were grown at 310 K in 650 ml LB medium containing 50 µg l^{-1} ampicillin. Protein expression was induced by the addition of arabinose to a final concentration of 0.2% and growth was continued at 298 K for 16 h. Cells were harvested by centrifugation at 6400*g* and resuspended in 50 m*M* 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES) pH 7.5, 500 m*M* NaCl and 10 m*M* imidazole (buffer *A*) prior to lysis by sonication.

The sonicated extract was clarified by centrifugation at 50 000*g* and applied at 277 K onto a 25 ml chelating Sepharose (GE Healthcare) affinity column charged with three column volumes of 50 m*M* nickel chloride in 50 m*M* sodium phosphate buffer pH 4 prior to equilibration in buffer *A*. After loading of the crude extract, the column was washed with ten column volumes of *A* followed by a further ten column volumes of buffer *A* supplemented with 100 m*M* imidazole pH 8.0. YgfE was then eluted from the column in buffer *A* containing 500 m*M* imidazole pH 8.0. The protein was dialysed into 20 m*M* HEPES pH 7.5 and concentrated to 12 mg ml^{-1}, as determined by BioRad protein assay.

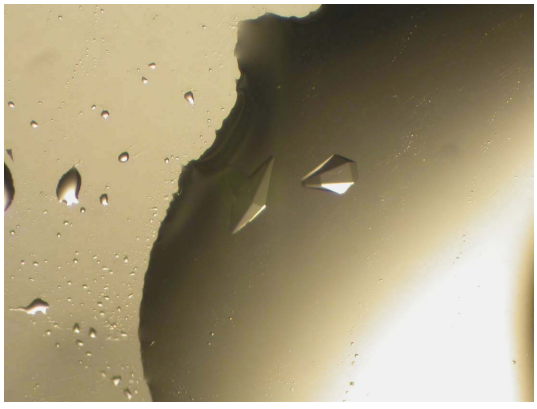


Figure 1
Crystals of *E. coli* YgfE obtained from 0.1 *M* HEPES pH 7.0, 18% PEG 6000, 5 m*M* $CaCl_2$. These crystals appear after within 2–3 d after incubation of the plate at 291 K and have a maximum size of 0.5 mm in any one dimension.

```
E. coli  YgfE          MS-AQPVDIQIFGRSLRVNCPPDQRDALNQAADDLNQRLQDLKERTRVTNT
P. aeruginosa ZapA     MSQSNTLTVQILDKEYCINCPDDERANLESAARYLDGKMREIRSSGKVIGADRVAVMAAI
                       ** ::.: :**:.:.  :*** *:*  *:.**  *: :::::::.

E. coli YgfE           NISYELAQEKAKTRDYAASMEQRIRMLQQTIEQALLEQGRITEKTNQNFE
P. aeruginosa ZapA     NITHDLLHRKERLDQESSSTRERVRELLDRVDRALANPADAGEA------ 104
                       **:::* :.* :  : ::* .:*:* * : :::** : .   *
```

Figure 2
CLUSTALW sequence alignment (http://www.ebi.ac.uk/clustalw/index.html) of *E. coli* K12-MG1655 YgfE (TIGR reference b2901) with *P. aeruginosa* PAO1 ZapA (TIGR reference PA5227).

Table 1
Data-collection and processing statistics.

Values in parentheses correspond to the outer resolution shell.

Wavelength (Å)	0.933
Space group	$P6_122$
Unit-cell parameters	
a (Å)	53.8
b (Å)	53.8
c (Å)	329.7
α (°)	90
β (°)	90
γ (°)	120
Matthews coefficient (Å^3 Da^{-1})	2.7
Molecules per AU	2
Solvent content (%)	55
Resolution range (Å)	30–1.8 (1.83–1.80)
Total observations	380518
Unique reflections	27357
Average $I/\sigma(I)$	47.9 (4.4)
R_{merge}	0.052 (0.242)
Completeness (%)	98.3 (78.2)

3. Crystallization and DLS assay

Prior to crystallization, samples of protein were filtered through a Whatman SpinX-centrifuge spin filter and subjected to dynamic light-scattering analysis using a DynaPro MS/X device to check for monodispersity. All crystallization experiments were performed using the hanging-drop vapour-diffusion method in a 24-well tissue-culture Linbro plate at 291 K. Initial crystallization trials were carried out using 0.5 ml reservoir solutions taken from the Hampton Research Crystal Screens I and II (Jancarik & Kim, 1991) or using 0.5 ml solutions taken from the Clear Strategy Screen buffered with 0.1 *M* MES pH 6.5 (Brzozowski & Walton, 2001). Drops consisting of 1 µl protein solution and 1 µl reservoir solution were used throughout. Crystals were obtained after several days using several conditions from the Clear Strategy Screen (CSS; Brzozowski & Walton, 2001) including CSS1 No. 1 (25% PEG 2K monomethylether, 0.3 *M* sodium acetate) and CSS1 No. 6 (0.1 *M* HEPES pH 7.0, 25% PEG 2K monomethylether, 0.8 *M* sodium formate). Additionally, crystals were obtained from a in-house screen developed for small proteins and peptides using conditions consisting of 0.1 *M* HEPES pH 7.0, 18% PEG 6000, 5 m*M* $CaCl_2$. These latter crystals were of a quality suitable for X-ray data collection, with dimensions in excess of 0.1 × 0.5 × 0.2 mm (Fig. 1).

3.1. X-ray crystallographic studies

Preliminary diffraction data were collected at 100 K in-house using an Enraf–Nonius Cu $K\alpha$ X-ray generator operating at 45 kV and 115 mA equipped with Osmic focusing mirrors and a MAR 345dtb image-plate detector. Complete data sets to the resolution quoted in this paper were collected at ESRF ID14.3. The crystals were vitrified in liquid nitrogen using 30%(*v*/*v*) glycerol in the mother liquor as cryoprotectant and were maintained at 100 K throughout data collection using an Oxford Cryostream system. Intensity data were indexed, integrated and scaled using the *HKL* programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Data-collection and processing statistics are given in Table 1.

During the course of this study, the structure of *P. aeruginosa* ZapA (PDB code 1w2e) became available. Although ZapA is a functional homologue of YgfE, there is a low level of sequence similarity between the two proteins (Fig. 2) and thus far we have been unable to obtain a molecular-replacement solution using ZapA as search model.

We are therefore investigating MIR and MAD strategies to obtain a solution to the structure of YgfE as part of a coordinated study investigating the interaction of YgfE, FtsZ and other cell-division proteins.

We are grateful for access and user support at the synchrotron facilities of the ESRF, Grenoble and SRS, Daresbury. We would like to thank Klaus Fütterer (University of Birmingham) for assistance with data collection. AB is funded by a BBSRC studentship; RPG is supported *via* the EPSRC-supported Molecular Organisation and Assembly in Cells (MOAC) doctoral training centre at Warwick, KAJ is supported by the Karl Trygger Foundation, SA and TD are supported by the Wellcome Trust and MRC Career Development fellowships, respectively.

References

Addinall, S. G. & Holland, B. (2002). *J. Mol. Biol.* **318**, 219–236.
Anderson, D. E., Gueiros-Filho, F. J. & Erickson, H. P. (2004). *J. Bacteriol.* **186**, 5775–5781.
Brzozowski, A. M. & Walton, J. (2001). *J. Appl. Cryst.* **34**, 97–101.
Feucht, A., Lucet, I., Yudkin, M. D. & Errington, J. (2001). *Mol. Microbiol.* **40**, 115–125.
Gueiros-Filho, F. J. & Losick, R. (2002). *Genes Dev.* **16**, 2544–2556.
Hale, C. A. & de Boer, P. A. (1997). *Cell*, **88**, 175–185.
Hale, C. A., Rhee, A. C. & de Boer, P. A. (2000). *J. Bacteriol.* **182**, 5153–5166.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Low, H. H., Moncrieffe, M. C. & Löwe, J. (2004). *J. Mol. Biol.* **341**, 839–852.
Lu, C., Stricker, J. & Erickson, H. P. (1998). *Cell Motil. Cytoskeleton*, **40**, 71–86.
Lutkenhaus, J. & Addinall, S. G. (1997). *Annu. Rev. Biochem.* **66**, 93–116.
Margolin, W. (2000). *FEMS Microbiol. Rev.* **24**, 531–548.
Moy, F. J., Glasfeld, E., Mosyak, L. & Powers, R. (2000). *Biochemistry*, **39**, 9146–9156.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
RayChaudhuri, D. (1999). *EMBO J.* **18**, 2372–2383.
Romberg, L. & Levin, P. A. (2003). *Annu. Rev. Microbiol.* **57**, 125–154.
Vaughan, S., Wickstead, B., Gull, K. & Addinall, S. G. (2004). *J. Mol. Evol.* **58**, 19–29.

The Binding of Single-Stranded DNA and PNA to Single-Walled Carbon Nanotubes Probed by Flow Linear Dichroism

Jascindra Rajendra and Alison Rodger*[a]

Abstract: The binding of single-stranded DNAs and a neutral DNA analogue (peptide nucleic acid, PNA) to single-walled carbon nanotubes in solution phase has been probed by absorbance spectroscopy and linear dichroism. The nanotubes are solubilised by aqueous sodium dodecyl sulfate, in which the nucleic acids also dissolve. The linear dichroism (LD) of the nanotubes, when subtracted from that due to the nanotubes/nucleic acid samples, gives the LD of the bound nucleic acid. The binding of the single-stranded DNA to the single-walled nanotubes is quite different from that previously observed for double-stranded DNA. It is likely that the nucleic acid bases lie flat on the nanotube surface with the backbone wrapping round the nanotube at an oblique angle in the region of 45°. The net effect is like beads on a string. The base orientation with the single-stranded PNA is inverted with respect to that of the single-stranded DNA, as shown by their oppositely signed LD signals.

Keywords: carbon · DNA · linear dichroism · nanotubes · PNA

Introduction

Carbon nanotubes, a comparatively recently recognised allotrope of carbon,[1–3] are remarkable materials that are presently being widely studied because of their electronic properties and range of potential applications as semiconductors, catalysts, optical devices and so forth.[4,5] Nanotubes are well-ordered hollow graphitic nanomaterials that vary in length from several hundred nanometers to several micrometers and have diameters of 0.4 to 2 nm, for single-walled carbon nanotubes (SWNT), and 2 to 100 nm, for coaxial multiple-walled carbon nanotubes (MWNT), depending on the number of concentric tubes.[6] A SWNT can be described as a single graphene sheet rolled into a tube, whereas a MWNT contains overlapping cylindrical tubes, like a coaxial cable. These tubes are proposed to be either metallic or semiconducting, depending on how the sheet is rolled up and its diameter. One of the attractions of SWNTs is that if they are fabricated with biomacromolecules (proteins, carbohydrates and nucleic acids)[7,6] attached to their surface, they have the potential of being used in antigen recognition, enzyme-catalysed reactions[8] and a wide range of DNA-hybridisation applications.[9]

The literature reports a range of molecules binding to SWNTs,[1–9] but as yet methods for characterizing the interaction, particularly in solution, are lacking. Most of the available literature data on DNA/SWNT systems studied by a wide range of techniques does not refer to the solution phase. A recent report of DNA solubilizing nanotubes described the use of AFM and absorbance spectroscopy to prove they were in solution; however, there was no attempt to characterise the interaction.[10] We have recently shown[11] that Couette flow linear dichroism (LD), the differential absorbance of light polarised parallel and perpendicular to a sample orientation direction, could be used to probe the binding of small aromatic molecules and double-stranded (ds) DNA to SWNTs. Interpretation of the flow LD data ideally requires the spectroscopy of the species being studied to be understood, and in particular their transition moments to be known. In the case of nanotubes, there is very little information in the literature about their spectroscopy in the near and far UV regions. In fact, it is usually assumed that there is none, as this region is dominated by light-scattering artifacts in most spectrometers. We showed[11] that the sample of CNTs (SWNTs formed by arc discharge and purified by refluxing with nitric acid) used in our work has an unstructured UV absorbance spectrum that gives rise to a large negative LD signal with maximum at 225 nm. The sign of this signal means that the dominant transition polarisa-

[a] J. Rajendra, Dr. A. Rodger
Department of Chemistry, University of Warwick
Coventry, CV4 7AL (UK)
Fax: (+44) 24-765-24112
E-mail: a.rodger@warwick.ac.uk

 DOI: 10.1002/chem.200500093

tion of these SWNTs lies at more than 54.7° from the nanotube long axis, since the reduced LD (LD^r) may be expressed as Equation (1)[12], in which α is the angle between the orientation axis and the transition moment of interest, S is the orientation parameter (1 for perfect orientation and 0 for an unoriented sample) and A is the sample's isotropic absorbance.

$$LD^r = \frac{LD}{A} = \frac{3}{2}S(3\cos^2\alpha - 1) \quad (1)$$

Assuming the SWNT has cylindrical symmetry about its long axis, ligand-binding geometry is best defined in terms of the normal to the cylinder surface. So, in principle the LD could yield the orientation of ligands on the SWNT by using Equation (2)[11,13,14] in which β is the angle the transition moment of interest makes with the normal to the nanotube surface.

$$LD^r = \frac{3S}{4}(1 - 3\cos^2\beta) \quad (2)$$

In practice, at this stage, we are not able to measure the degree of orientation of the SWNT, so our geometric conclusions are qualitative rather than quantitative. An additional application of the LD data was suggested by the fact that spectra of both anthracene and naphthalene were significantly perturbed by the SWNT, and the intensity of the SWNT's signal was also changed. This suggests that LD may also be a means of probing the type of interaction anchoring a ligand onto the SWNT. Before this is really feasible, we need to understand more about the interaction of ligands with SWNTs. In this paper, we use LD to investigate the interaction of a range of different single-stranded (ss) nucleic acids binding to sodium dodecyl sulfate solubilised SWNTs. Literature reports suggest that ss DNA binds better to nanotubes than does ds DNA[15] (ds = double-stranded). A recent report has concluded from surface-enhanced infrared absorption (SEIRA) studies that ss DNA wraps SWNT more efficiently than ds DNA,[16] and the proposed binding geometry may be the reason for this. So, we anticipated that LD could be used to probe the role played by the structure of the ss DNA upon interaction with SWNTs.

Results and Discussion

The DNA absorption spectrum is the net effect of overlapping $\pi \rightarrow \pi^*$ transitions of the purine and pyrimidine bases. These transitions are polarised in the plane of the bases, and when polymeric ds DNA is flow-oriented, their polarisations are all approximately perpendicular to the DNA helix axis.[12] This results in ds DNA having a negative LD signal under its absorbance bands. We showed previously that extensively sonicated double-stranded calf thymus (ct) DNA has no detectable intrinsic LD signal, though it does acquire a small negative LD signal at ~260 nm (the absorbance maximum) when in the presence of SWNTs.[11] To our surprise we found that all the supposedly single-stranded DNAs used in this work have intrinsic LD signals. By way of contrast, duplex DNAs of 250 base pairs scarcely orient.[17] In each case, the ss DNA LD maximum was negative and at the same wavelength as the absorbance maximum. This means that the ss DNAs form long structures in solution.

Preparing solutions of SWNTs in the presence of DNA proved not to be straightforward. When ds DNA was melted in the presence of SWNTs, the sample came out of solution. When sonicated, pre-melted and cooled ds DNA (which should be ss DNA) was added to a solution of SWNTs, everything remained in solution. However, as noted above, the isolated DNA has an intrinsic LD signal (of $\sim 10^{-4}$) even though it had no detectable signal before melting. The effect was much the same whether the DNA was fast cooled (by putting it onto ice) or slow cooled (leaving it to cool at 1 °C min^{-1} in a spectrometer). Single-stranded ct-DNA obtained from Sigma had the same feature. Despite these issues, we believe we are seeing the LD of single-stranded DNA bound to SWNTs in Figure 1c, as the LD difference spectra obtained from subtracting the SWNT LD spectrum from DNA/SWNT LD spectra (Figure 1) has a negative peak at ~275 nm and a positive peak at ~230 nm, rather than a single negative peak at the absorbance maximum near 260 nm.

To understand why the bisignate LD spectrum arises, it is necessary to consider the spectra of the DNA bases, of which each has more than one transition in the region of interest (Figure 2a, b). The fact that the 230 nm region of the spectrum gives a positive LD signal means [following Eq. (2)] that β (the angle between the normal to the nanotube and the transition moment) for these transitions is greater than 55°. In other words, on average the transitions are within 35° of the nanotube long axis. Figure 2c is a schematic diagram indicating the orientations of the base transitions when the bases are linked to a simple backbone and the backbone is wrapped round the nanotube at an inclination angle of ~45°. The 230 nm region transitions of C[18] and A[24] are within 10° of the nanotube axis; this orientation would give a positive LD signal. The 244 and 200 nm transitions of G,[19] which will have some intensity at 230 nm, would also contribute a small positive signal, as oriented in Figure 2c. The T's contribution would be small and negative.[20] The Figure 2c geometry would also result in all bases giving a negative contribution in the ~270 nm region. Although this analysis is by no means definitive, it suggests an orientation that is consistent with the data. The analysis would be improved if it were possible to calculate the LD^r for the spectrum. This is currently hindered by the quality of the absorbance data and the significant contribution made by scattering (which does not follow a simple wavelength–power law).[21,22] It should be noted, however, that the situation in reality is almost certainly a mixture of binding modes. Something like the one proposed may be the dominant one. To avoid the complications introduced by having the transitions of four bases being represented in the spec-

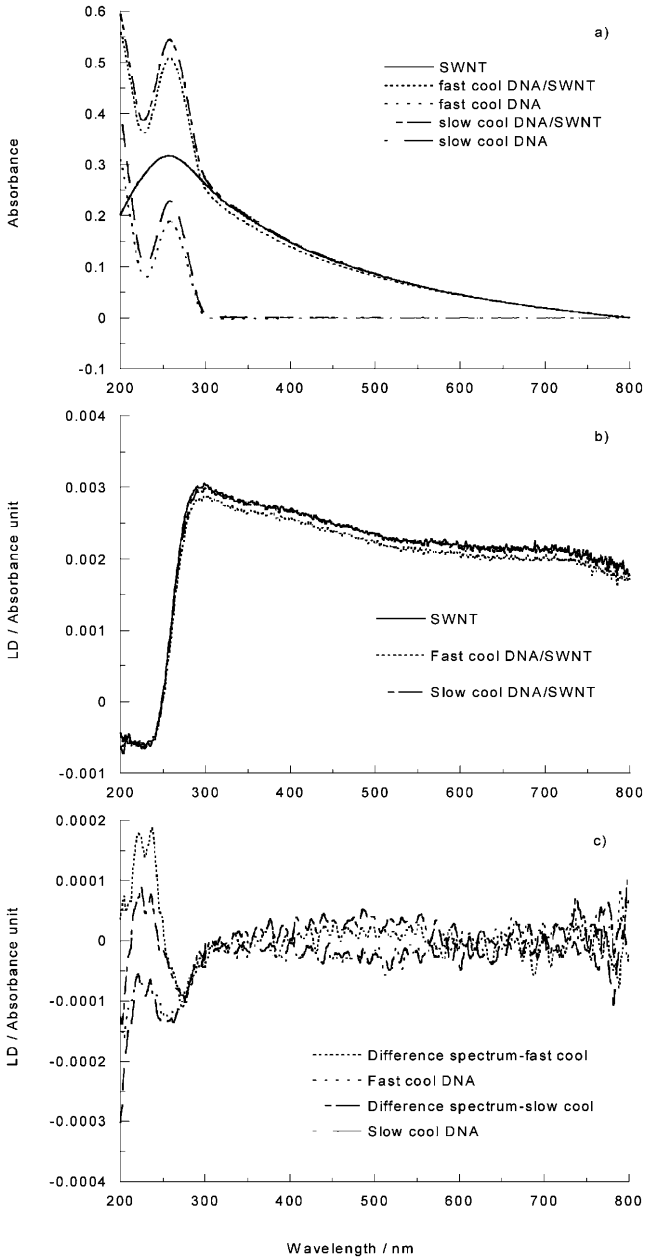


Figure 1. a) Absorbance spectra of SWNT (0.1 mg mL^{-1}), fast-cooled ds DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex, fast-cooled ds DNA, slow cooled ds DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex in SDS (9 mM) and slow cooled ds DNA. b) LD spectra of SWNT (0.1 mg mL^{-1}), fast-cooled DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex and slow-cooled DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex in SDS (9 mM). c) LD spectra of fast-cooled and slow-cooled DNA in SDS (9 mM) and difference LD spectra of fast-cooled DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) minus SWNT (0.1 mg mL^{-1}) and slow-cooled DNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) minus SWNT (0.1 mg mL^{-1}) complex, all in (9 mM) SDS. All spectra have SDS (9 mM) baselines subtracted.

trum, we also considered synthetic DNAs of well-defined sequence, as discussed below.

From Figure 1a it is apparent that the fast-cooled (though not slow-cooled) DNA assists the solubility of SWNT in

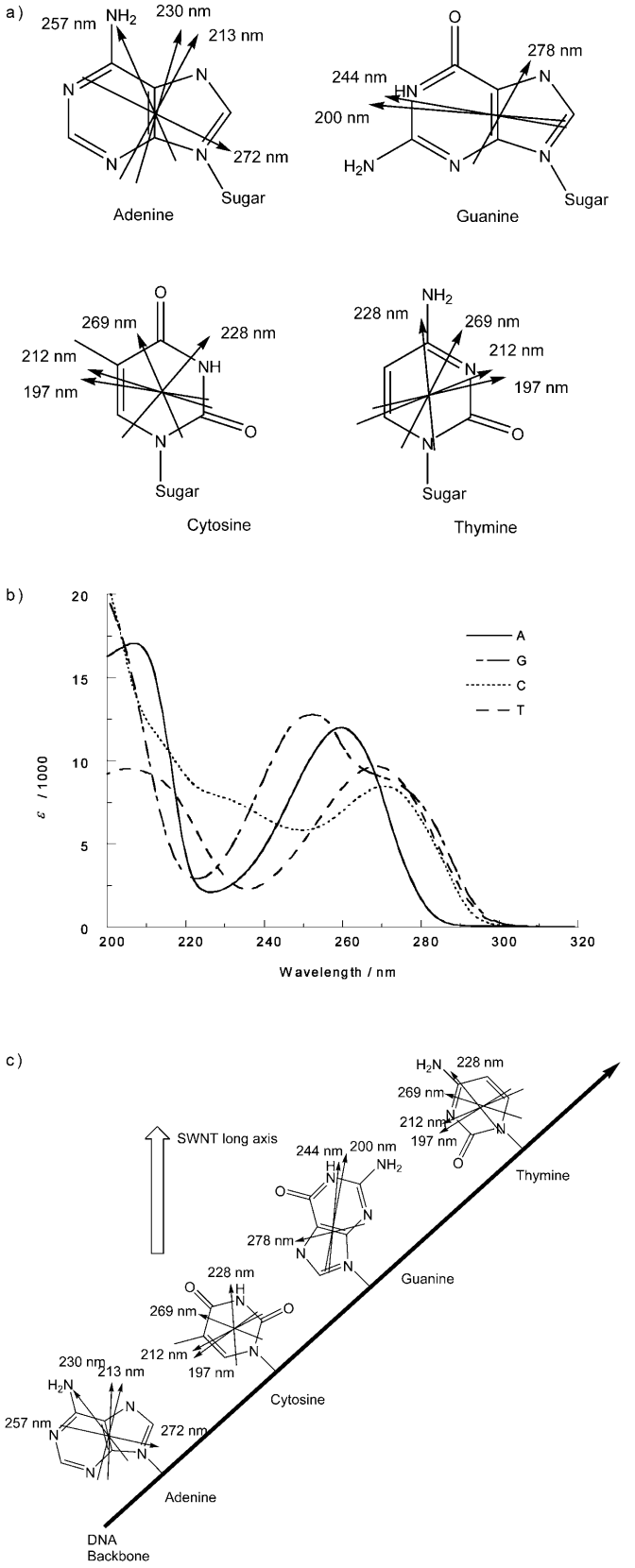


Figure 2. a) Probable transition polarisations for UV transitions of adenine, guanine, cytosine and thymine. b) UV spectra of the DNA nucleotides. c) Schematic diagram of orientation of DNA bases on a SWNT consistent with the observed LD spectra. For PNA the SWNT is rotated 90°.

sodium dodecyl sulfate (SDS), as the light scattering of the spectra reduced when fast-cooled DNA is added to the SWNT solution. This DNA also gives a larger low-wavelength DNA LD signal, suggesting its bases prefer lying flat on the SWNT following the above analysis. The fast-cooled DNA absorbance is smaller than that of the same concentration of slow-cooled DNA, suggesting it has more π–π stacking. Rather than implying the fast-cooled DNA has more ds DNA (the usual interpretation of such a hypochromic effect), it may be that the fast-cooled sample is less tangled and so better stacked within its strands. This would also make it more available than the slow-cooled DNA for binding in the flat-on-surface mode. Molecular modelling has suggested that ss DNA can adopt many different modes of binding to SWNTs, including helical wrapping with different pitches,[23] but it is not possible to differentiate them with LD at this stage. The spectra for the Sigma ss DNA (data not shown) are very similar to those for the fast-cooled DNA.

Poly(deoxyadenylate): Poly(dA) is the single nucleotide ss DNA least likely to form hydrogen bonds, though as is evident from Figure 3c it still has an intrinsic LD signal with a negative maximum at 255 nm. When it is in solution with SWNT, this shifts to 260 nm. The DNA–SWNT complex also has a small positive signal at ~225 nm (Figure 3c). Adenine has strong transitions centred at 257 and 213 nm and two weaker transitions at 230 and 272 nm (Figure 2).[24] The SWNT binding geometry proposed for ss DNA in Figure 2c would give rise to the observed LD spectrum. As is the case with ss ct-DNA (both melted and from Sigma), the light scattering is reduced in the poly(dA)/SWNT complex spectra when compared to SWNT spectra (Figure 3). This indicates that the poly(dA) facilitates the solubility of SWNT in SDS.

DNA oligomers: To our surprise, even a random base hexamer $d(N)_6$ 5′-phosphate gave an intrinsic negative LD signal at ~260 nm of magnitude −0.00025 (data not shown). Therefore, we were forced to conclude that even these small fragments link to form long strands. There was no evidence in the LD spectra (data not shown), however, that this hexamer interacted with the SWNT. By way of contrast, the octamer $(dCdT)_4$ ($T_m = 17\,°C$) gave almost no intrinsic LD signal (Figure 4c), but in the presence of SDS-solubilised SWNTs an LD signal was observed (Figure 4c). The light scattering of this sample is higher (Figure 4a) than that of the SWNT (see below), in contrast to the situation for the polymeric DNAs, in which the SWNT light scattering is reduced upon DNA addition. Despite this, we can clearly see a negative LD band at ~267 nm and a positive band at ~238 nm for the difference spectrum of $(dCdT)_4$/SWNT−SWNT. The LD signs again are consistent with the oligomer wrapping around SWNT in a tilted fashion, as illustrated in Figure 2c. Zheng et al. saw no evidence for this oligomer binding to SWNTs, suggesting that the LD may be a more sensitive test than absorbance spectroscopy.[15,23]

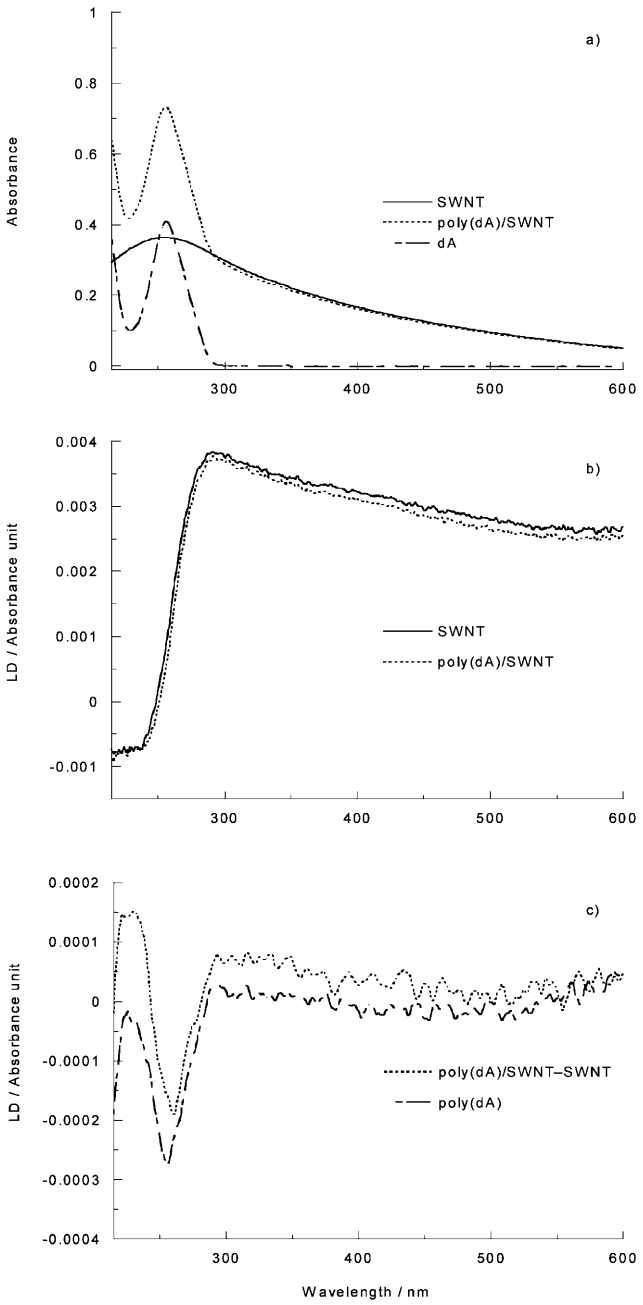


Figure 3. a) Absorbance spectra of SWNT (0.1 mg mL^{-1}), poly(dA) (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex and poly(dA) (0.1 mg mL^{-1}). b) LD spectra of SWNT (0.1 mg mL^{-1}) and poly(dA) (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex in SDS (9 mM). c) LD spectrum of d(A) (0.1 mg mL^{-1}) in SDS (9 mM) and LD difference spectrum of poly(dA) (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) minus SWNT (0.1 mg mL^{-1}). All spectra have SDS (9 mM) baselines subtracted.

Peptide nucleic acid (PNA): We thought that the anionic SDS that seemed to be required to get the SWNTs into solution would hinder the anionic DNA from binding to the nanotubes. Thus, we investigated the interaction between SWNTs and $(CT)_4$ peptide nucleic acid (PNA), which is a neutral DNA analogue with a peptide backbone composed of *N*-(2-aminoethyl)glycine units rather than a sugar phos-

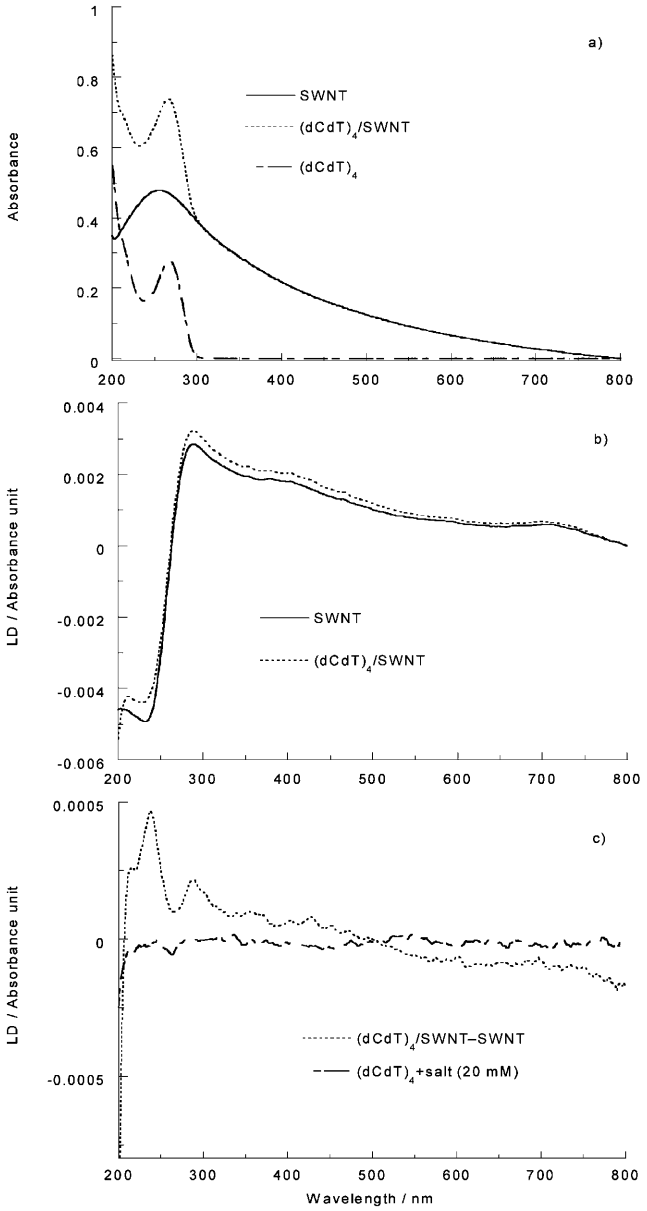


Figure 4. a) Absorbance spectra of SWNT (0.1 mg mL^{-1})/salt (20 mM) in SDS (9 mM), $(dCdT)_4$ (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1})/salt (20 mM) complex in SDS (9 mM) and $(dCdT)_4$/salt absorbance of the same concentration. b) LD spectrum of SWNT (0.1 mg mL^{-1})/salt (20 mM) compared with the $(dCdT)_4$ (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1})/salt (20 mM) spectrum. c) LD spectrum of $(dCdT)_4$ (0.1 mg mL^{-1})/salt (20 mM) spectrum and difference LD spectrum of $(dCdT)_4$ (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1})/salt (20 mM) complex minus SWNT (0.1 mg mL^{-1})/salt (20 mM) zeroed at 500 nm. All spectra had SDS (9 mM) baselines subtracted.

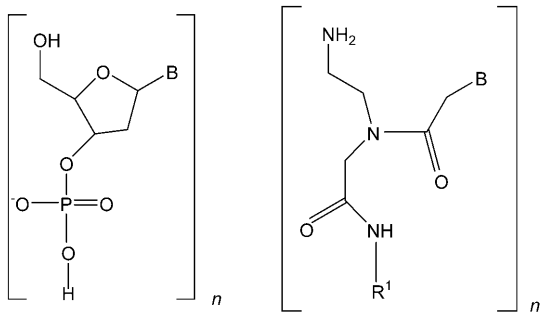


Figure 5. Chemical structures of PNA (left) and DNA (right) backbones. B = base.

phate backbone (Figure 5). PNA is capable of sequence-specific recognition of DNA and RNA obeying the Watson–Crick hydrogen-bonding rules, so it could bind to SWNTs with the same geometry as DNA.[25]

The spectra for SWNT and $(CT)_4$ PNA are shown in Figure 6. As with the same sequence DNA oligomer, the light scattering is higher (and somewhat different in shape) for the PNA/SWNT complex than for the SWNT. This contrasts with the situation for the DNA polymers, in which it appeared they helped to solubilise the DNA, reducing the scattering. Whether this difference relates to the way in which the SDS is displaced by the longer and shorter nucleic acids is not clear at this stage. It should be noted that there is a net scattering "baseline" left in the spectra of Figures 4 and 6, due to the different shape of the scattering profiles of the SWNT with and without the oligomers. It may in any case be concluded that PNA is adsorbed on to the SWNT. The PNA itself only has a very small (negative) LD signal at 260 nm. The difference LD spectrum of PNA/SWNT (Figure 6c), like the DNA/SWNT system, has two bands, one at ~290 nm and the other at ~240 nm. These are slightly red-shifted compared with the DNA analogue. More intriguingly, the sign pattern of the PNA system is inverted relative to that of DNA. Thus, whatever the binding mode of DNA on the SWNT, PNA is oriented quite differently. The PNA backbone has even more flexibility than the DNA backbone, and the LD signs would suggest it wraps around the SWNT so that the DNA bases lie approximately perpendicular to those illustrated in Figure 2c. A means of achieving this would be for it to wrap with the opposite sense, which can be visualised by inserting the SWNT with its axis horizontal rather than vertical in Figure 2c Alternatively, each base in Figure 2c can be rotated 180° about the link from the base to the backbone. Although the PNA backbone can be aligned with that of DNA, it is intrinsically flexible and not itself chiral. So the lower energy arrangements for PNA on the SWNT could well be different from those of the DNA.

Conclusions

In this paper, we have established the potential of flow linear dichroism to probe the binding of single-stranded DNAs to SWNTs. DNAs varying in size from a few hundred bases to short oligomers were investigated, and a range of different sequences were found to interact. Although only qualitative analysis of the data is possible at this stage, one could envisage determining the orientation of the DNA

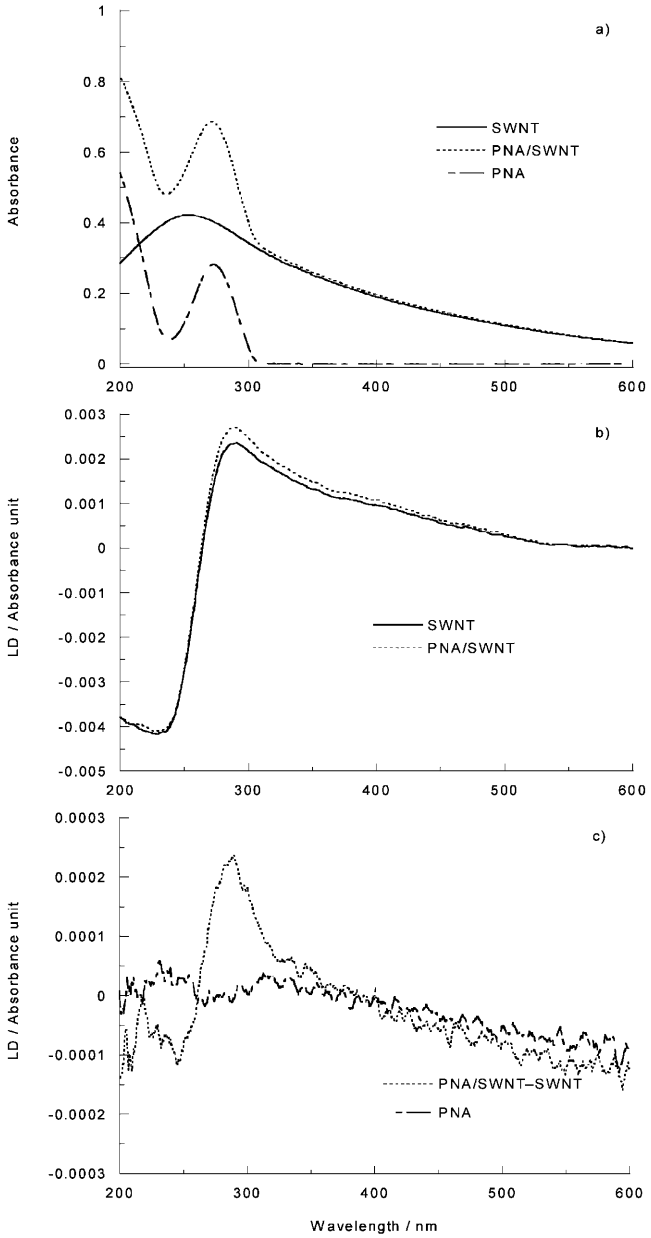


Figure 6. a) Absorbance spectra of SWNT (0.1 mg mL^{-1}) in SDS (9 mM), PNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex in SDS (9 mM) and PNA (0.1 mg mL^{-1}) in SDS (9 mM). b) LD spectra of SWNT (0.1 mg mL^{-1}) in SDS and PNA (0.1 mg mL^{-1})/SWNT (0.1 mg mL^{-1}) complex in SDS (9 mM). c) LD spectrum of PNA (0.1 mg mL^{-1}) in SDS (9 mM) and difference spectrum of PNA/SWNT minus SWNT, zeroed at 400 nm. All spectra have SDS (9 mM) baselines subtracted.

bases on the SWNT by deconvoluting transition polarisations of the bases. A geometry consistent with the observed data is that in which the DNA backbone wraps round the DNA at an oblique angle and the bases lie flat on the nanotube surface. We presume this involves displacing the anionic SDS from the SWNT surface.

The LD signals for the single-stranded DNAs were not as large as those previously observed[11] for sonicated double-stranded ct-DNA. This might be because of the cooperative effect of a large number of small interactions coupled with the intrinsically greater rigidity of the duplex DNA. In contrast to the situation with ds DNA, in which a single negative band at the absorbance maximum was observed, the spectrum observed in each case for ss DNA is a couplet of bands with a negative one at ~280 nm and a positive one at ~230 nm. The polymeric DNAs also reduced the light scattering observed for the SWNTs, whereas the oligomers increased it, suggesting the latter increase the SWNT size and the former help solubilise it. The neutral DNA analogue, $(CT)_4$ peptide nucleic acid, intriguingly gave the opposite signed spectrum from that of the DNA $(dCdT)_4$, suggesting an orientation of the PNA on SWNTs in which the bases are rotated ~90° with respect to that observed for DNA.

Experimental Section

Materials: SWNTs (synthesised by the catalytic arc discharge method) were obtained from Dynamic Enterprises, were purified by refluxing in 3 M nitric acid at 120 °C for 13.5 h and were then washed with water (18.2 MΩ).[11] Purity was investigated by transmission electron microscopy (TEM) as described previously.[11] In order to overcome problems due to inhomogeneity of the samples, stock solutions of SWNT (0.5 mg mL^{-1}) were prepared. The stock solution of SWNTs was obtained by sonicating the SWNTs for 2 min in aqueous sodium dodecyl sulfate (SDS, Sigma, 9 mM) at a concentration slightly above the critical micelle concentration (which is 8.5 mM) of SDS to give a viscous solution of concentration 0.5 mg SWNT per mL of SDS.[26] The final concentration of SWNT was maintained at no more than 0.1 mg mL^{-1} in all samples in order to avoid excessive absorption. The DNAs were then added to the SWNT SDS solutions as outlined below.

Double- and single-stranded calf thymus DNA (ct-DNA) were obtained from Sigma–Aldrich Chemical Company. Poly(dA) and the random DNA hexamer $(pd(N)_6)$ were obtained from Amersham Biosciences and the DNA and PNA oligomers with the sequence $(CT)_4$ were obtained from Sigma Genosis and Eurogentec, respectively. Stock solutions of the polymeric DNAs were prepared in SDS (9 mM); stock solutions of the oligomers were prepared in water (18.2 MΩ). All spectroscopy experiments were performed with aqueous SDS (9 mM) as the solvent. For the spectroscopy experiments, an aliquot of DNA was added to a SWNT suspension (~0.1 mg mL^{-1}) in SDS to a concentration of 0.1 mg mL^{-1}, though in some cases further dilution (with SDS; 9 mM) was required to avoid excessive absorbance. Short ds ct-DNA of ~200–400 base pairs (as determined by gel electrophoresis[11]) was obtained by sonicating an aqueous SDS solution of ct-DNA for about 2–3 h.[27] All preparations were left overnight to equilibrate before spectroscopic measurements were performed; in fact, the DNA/SWNT samples gave the same results if the spectra were measured immediately upon mixing. Samples left standing for a few days, however, showed some precipitation from solution and gave unsatisfactory spectra. Though SDS was used as baselines for the data presented below, we were also able to use samples without rotation for the LD spectra as they overlaid very well. In fact, this coincidence of the baselines was a good check on system performance.

Spectroscopy

Absorbance: UV-visible absorbance spectra were recorded using a Cary 1E spectrophotometer.

Linear dichroism (LD): LD spectra were recorded using a Jasco J-715 circular dichroism spectropolarimeter with extended sample compartment adapted for LD measurements. In LD spectra the difference in anisotropic absorption of light polarised in planes parallel ($A_{\parallel}$) and perpendicular ($A_{\perp}$) to the direction of orientation[12] [Eq. (3)] is measured.

$$LD = A_{\parallel} - A_{\perp} \quad (3)$$

The Couette cell used to flow orient the samples was a small-volume quartz LD cell, the outer quartz cylinder of which had an internal diameter of ~3 mm and rotated, while the inner cylinder was a stationary ~2.5 mm diameter quartz rod.[28] This cell was equipped with focusing lenses before and after the sample; this minimised light scattering. Its demountable sample holding capillary facilitated cleaning. The voltage applied to the cell in all experiments was 3.0 V, which corresponds to a rotation speed of ~1000 rpm.

[1] Q. Liang, L. Z. Gao, Q. Li, S. H. Tang, B. C. Liu, Z. L. Yu, *Carbon* **2001**, *39*, 897–903.
[2] T. Saito, K. Matsushige, K. Tanaka, *Phys. B* **2002**, *323*, 280–283.
[3] P. Bernier, W. Maser, C. Jouret, A. Loiseau, M. L. de la Chapelle, S. Lefrant, R. Lee, J. E. Fischer, *Carbon* **1998**, *36*, 675–680.
[4] E. Kymakis, I. Alexandou, G. A. J. Amaratunga, *Synth. Met.* **2002**, *127*, 59–62.
[5] E. K. Hobbie, H. Wang, H. Kim, C. C. Han, E. A. Grulke, J. Obrzut, *Rev. Sci. Instrum.* **2003**, *74*, 1244–1250.
[6] Y. Lin, S. Taylor, H. Li, K. A. S. Fernando, L. Qu, W. Wang, L. Gu, B. Zhou, Y. Sun, *J. Mater. Chem.* **2004**, 14, 527–541.
[7] D. L. Nelson, M. M. Cox, *Principles of Biochemistry*, 3rd ed., Worth Publishers, **1999**.
[8] K. Rege, N. R. Raravikar, D. Y. Kim, L. S. Schadler, P. M. Ajayan, J. S. Dordick, *Nano Lett.* **2003**, *3*, 829–832.
[9] S. E. Baker, W. Cai, T. L. Lasseter, K. P. Weidkamp, R. J Hamers, *Nano Lett.* **2002**, *2*, 1413–1417.
[10] N. Nakashima, S. Okuzono, H. Murakami, T. Nakai, K. Yoshikawa, *Chem. Lett.* **2003**, 32, 456–457.
[11] J. Rajendra, M. Baxendale, L. G. Dit Rap, A. Rodger, *J. Am. Chem. Soc.* **2004**, *126*, 11182–11188.
[12] A. Rodger, B. Nordén, *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, **1997**.
[13] A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. T. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez, M. J. Hannon, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4051–4057.
[14] M. Ardhammar, P. Lincoln, B. Nordén, *J. Phys. Chem. B* **2001**, *105*, 11363–11368.
[15] M. Zheng, A. Jagota, E. D. Semke, B. A. Diner, R. S. Mclean, S. R. Lustig, R. E. Richardson, G. T. Nancy, *Nat. Mater.* **2003**, 2, 338–342.
[16] G. I. Dovbeshko, O. P. Repnytska, E. D. Obraztsova, Y. V. Shtogun, *Chem. Phys. Lett.* **2003**, 372, 432–437.
[17] A. Rodger, D. J. Halsall, unpublished results **2003**.
[18] F. Žaloudek, J. S. Novros, L. B. Clark, *J. Am. Chem. Soc.* **1985**, *107*, 7344–7351.
[19] L. B. Clark, *J. Am. Chem. Soc.* **1977**, *99*, 3934–3938.
[20] A. L. Williams, Jr., C. Cheong, I. Tinoco, Jr., L. B. Clark, *Nucleic Acids Res.* **1986**, *14*, 6649–6659.
[21] J. Nordh, J. Deinum, B. Nordén, *Eur. Biophys. J.* **1986**, *14*, 113–122.
[22] R. Marrington, M. Seymour, A. Rodger, polymerisation of tubulin studied by linear dichroism, unpublished results.
[23] M. Zheng, A. Jagota, M. S. Strano, A. P. Santos, P. Barone, G. S. Chou, B. A. Diner, M. S. Dresselhaus, R. S. Mclean, G. B. Onoa, G. G. Samsonidze, E. D. Semke, M. Usrey, M. D. J. Walls, *Science* **2003**, *302*, 1545–1548.
[24] A. Holmén, A. Broo, B. Albinsson, B. Nordén, *J. Am. Chem. Soc.* **1997**, 119, 12240–12250.
[25] F. Hook, A. Ray, B. Nordén, B. Kasemo, *Langmuir* **2001**, *17*, 8305–8312.
[26] S. K. Doorn, R. E. Fields, H. Hu, M. A. Hamon, R. C. Haddon, J. P. Selegue, V. Majidi, *J. Am. Chem. Soc.* **2002**, *124*, 3169–3174.
[27] H. I. Elsner, E. B. Lindblad, *DNA* **1989**, *8*, 697–701.
[28] R. Marrington, T. R. Dafforn, D. J. Halsall, R. Mortimer, T. Andrews, A. Rodger, *Biophys. J.* **2004**, *87*, 2002–2012.

Received: January 26, 2005
Published online: June 14, 2005

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Chemical Composition and Antibacterial Activity of the Essential Oil and the Gum of *Pistacia lentiscus* Var. *chia*

CHRISTINA KOUTSOUDAKI,[†] MARTIN KRSEK,[‡] AND ALISON RODGER*,[†]

Department of Chemistry and Department of Biological Sciences, University of Warwick, Warwick, Coventry CV4 7AL, United Kingdom

The essential oil and gum of *Pistacia lentiscus* var. *chia*, commonly known as the mastic tree, are natural antimicrobial agents that have found extensive uses in medicine in recent years. In this work, the chemical composition of mastic oil and gum was studied by GC−MS, and the majority of their components was identified. α-Pinene, β-myrcene, β-pinene, limonene, and β-caryophyllene were found to be the major components. The antibacterial activity of 12 components of mastic oil and the oil itself was evaluated using the disk diffusion method. Furthermore, attempts were made to separate the essential oil into different fractions in order to have a better picture of the components responsible for its antibacterial activity. Several trace components that appear to contribute significantly to the antibacterial activity of mastic oil have been identified: verbenone, α-terpineol, and linalool. The sensitivity to these compounds was different for different bacteria tested (*Escherichia coli*, *Staphylococcus aureus*, and *Bacillus subtilis*), which suggests that the antibacterial efficacy of mastic oil is due to a number of its components working synergistically. The establishment of a correlation between the antibacterial activity of mastic oil and its components was the main purpose of this research. Mastic gum was also examined, but it proved to be more difficult to handle compared to the essential oil.

KEYWORDS: Mastic oil; antibacterial activity; GC−MS; disk diffusion

INTRODUCTION

Mastic is a white, semitransparent, natural resin that is obtained as a trunk exudate from mastic trees. The mastic tree is an evergreen bush that thrives in the Eastern Mediterranean area, but only in the southern part of the island of Chios, a Greek island, does the plant produce resin that congeals. Its scientific name is *Pistacia lentiscus*, of the Anacardiaceae family. Mastic gum has numerous qualities and uses and is now exported to many countries. In medicine, a lot of research has been undertaken on the properties of mastic gum. For example, mastic gum has been used in clinical trials on patients with peptic ulcers (*1*). The administration of mastic (1 g daily) relieved the pain and healed the stomach and duodenal ulceration in the majority of the patients within 2 weeks. The same group of researchers (*2*) confirmed that mastic gum kills *Helicobacter pylori*, at concentrations as low as 0.06 mg/mL. In an earlier study (*3*), the effect of mastic has been studied on experimentally induced gastric and duodenal ulcers in rats. Mastic at an oral dose of 500 mg/kg produced a significant reduction of gastric secretions, protected cells, and reduced the intensity of gastric mucosal damage. The in vitro antimicrobial activity of *P. lentiscus* extracts has also been tested on bacteria and fungi (*4*). Of the different plant extractions (decoctions, infusions, macerations, and extracts from petroleum ether and from ethanol), decoctions showed the best antibacterial activity. It has also been found that chewing mastic gum prevents plaque formation or reduces it when it has already been formed on those teeth surfaces that can be reached by the mass of mastiche during its methodical chewing (*5*).

In surgery, byproducts of mastic gum are used for the production of special stitches that are eventually absorbed by the human body. In dentistry, mastic acts as an oral antiseptic and tightens the gums (*5*), and for that reason it is used in toothpastes and chewing gums. The essential oil of mastic gum is also used in perfumery and in the cosmetic industry (creams and other facial products) (*6*). Moreover, there are culinary uses of mastic, for example, in biscuits, ice cream, and mastic "sweets of the spoon".

The chemical composition of the mastic oil and mastic gum has recently been studied (*7*, *8*), but as yet no correlation between the antibacterial activity and the composition of mastic has been reported for the variety *chia*. The purpose of this study was to examine the chemical composition of this oil and gum and to examine the antibacterial activity of 12 individual components of mastic oil against three test organisms.

MATERIALS AND METHODS

Mastic oil (100% pure) and mastic gum, both of the harvest of 2002, were kindly provided by Chios Gum Mastic Growers Association

* Corrrsponding author. Telephone: +44(0)-24-76574696/76523234. Fax: +44(0)-24-76575795. E-mail: a.rodger@warwick.ac.uk.
[†] Department of Chemistry.
[‡] Department of Biological Sciences.

10.1021/jf050639s CCC: $30.25
Published on Web 09/02/2005

(Chios, Greece) and VIORYL S. A. (Athens, Greece). Five oil samples were used in the analysis: mastic oil as received, three successive fractions from distillation (see below), and the fraction of the oil that remained in the flask after the distillation. The oil was diluted in ethanol (50% v/v) in order to reduce the concentration of the pure oil and identify even its trace components. The distillate fractions (see below) were analyzed as collected. Mastic gum was ground with the help of a pestle and mortar and was then partially dissolved in ethanol (30 mg/mL), and the undissolved part of the gum was removed by filtration and disgarded. The ethanol soluble part of the gum was analyzed by GC−MS and GC−FID and was also tested for its antibacterial activity. Standards were also analyzed and tested for their antibacterial activity: α-pinene, β-myrcene, *p*-cymene, β-caryophyllene, verbenone, α-terpineol, methyl isoeugenol, limonene, β-pinene, linalool, γ-terpinene, and *trans*-anethole. All of the above components were of the highest purity available (above 97%) and were obtained from Sigma-Aldrich Chemical Co. (Dorset, UK), with the exception of α-terpineol and *trans*-anethole, which were obtained from Fisher Scientific (Leicestershire, UK). A 1% v/v solution in ethanol was prepared for each standard in order to be analyzed by GC−MS and GC−FID.

Distillation of Mastic Oil. A microdistillation setup was used to separate the mastic oil (2 mL) into fractions of differing volatilities. The sample was heated by an oil bath and the pressure in the distillation system was reduced to 20 mmHg. The first fraction was collected without any heating (maximum temperature 20 °C). When the distillation stopped, the mastic oil that remained in the flask was gradually heated to 21 °C and the distillate collected (fraction 2) until distillation ceased. No further distillate was collected, despite increasing the temperature of the oil bath to 140 °C. When the apparatus was removed from the oil bath, it was observed that the remaining essential oil had become very viscous and its color had changed from very pale yellow to very intense yellow. The flask was rinsed twice with ethanol. The first rinse became fraction 3 and the second fraction 4.

Chemical Composition of Mastic Oil and Gum. The GC−MS analysis of the samples was undertaken using a Shimadzu GC-17A, QP-5000 GC−MS system, operating in electron ionization (EI) mode with an ionization energy of 70 eV. The instrument was equipped with a Supelco SPB TM-1 capillary column (30 m, 0.25 mm i.d., 0.25 μm film thickness) with helium as carrier gas at 0.7 mL/min flow rate. Column temperature was initially kept for 1 min at 60 °C, gradually increased to 180 °C at a rate of 3.5 °C/min, and finally increased to 280 °C at a rate of 20 °C/min and kept there for 2 min. The injector and interface were set at 220 and 250 °C, respectively. The gas chromatograph operated in the split mode with a split ratio of 93:1. The mass spectrum was monitored starting at *m*/*z* 60 and ending at *m*/*z* 350, with a scan interval of 0.5 and a threshold of 400, and the solvent cut was set at 4 min. The injection volume was 1 μL. The injected solutions were (i) solution of mastic oil in ethanol (50% v/v), (ii) ethanol solutions of each standard (1% v/v), (iii) the ethanol-soluble part of mastic gum (30 mg/mL original sample before filtration), and (iv) the collected fractions from the distillation (see Distillation of Mastic Oil).

The chemical composition of mastic oil and gum was also analyzed using GC−FID, since the flame ionization detector is known to have higher sensitivity and the signal magnitude is to a reasonable approximation proportional to the analyte concentration, independent of its identity. The samples were prepared as for the GC−MS analysis. The GC−FID analysis of the samples was undertaken with a Shimadzu GC-17A, system, equipped with a capillary column SGE-BPX5 (30 m, 0.32 mm i.d., 0.5 μm film thickness). The carrier gas used was helium, the makeup gas was nitrogen, while hydrogen and air were used as ignition gases for the detector. The data system used was Shimadzu Class VP Chromatography Software. The method used was similar to that of GC−MS. Column temperature was initially kept for 1 min at 60 °C, gradually increased to 180 °C at a rate of 3.5 °C/min, and finally increased to 280 °C at a rate of 20 °C/min and kept there for 2 min. The injector and detector were set at 280 and 340 °C, respectively. The flow rate of the carrier gas in the column was kept constant at 1.0 mL/min, and the gas chromatograph operated in the split mode with a split ratio of 100:1. One microliter of the samples was injected manually, as for the GC−MS analysis.

Antibacterial Activity of Mastic Oil, Its Components, and Its Fractions. Three representative bacteria were selected for this study: *Escherichia coli* (Gram-negative rod), *Staphylococcus aureus* (Gram-positive cocci), and *Bacillus subtilis* (Gram-positive rod). The 12 individual components of mastic oil tested were α-pinene, β-myrcene, *p*-cymene, β-caryophyllene, verbenone, α-terpineol, methyl isoeugenol, limonene, β-pinene, linalool, γ-terpinene, and *trans*-anethole. All of the above components were of the highest purity available (above 97%) and were obtained from Sigma-Aldrich Chemical Co. (Dorset, UK), with the exception of α-terpineol and *trans*-anethole, which were obtained from Fisher Scientific (Fisher Scientific, Leicestershire, UK).

The disk diffusion susceptibility method (*11*−*13*) was used in order to examine the sensitivity of the bacteria of interest toward mastic oil, its components, and its collected fractions. Essential oils and many of their components have limited solubility in aqueous media, and this property was expected to cause difficulty in susceptibility test methods. To overcome this difficulty, a modified disk diffusion method was also tried. In preliminary experiments to examine the effect of solubilizing agent on the diffusion of components through the agar, two sets of media were prepared: one without the presence of detergent and one with the presence of Tween 80 detergent. Mueller−Hinton (MH) Agar was used, prepared according to the instructions of the manufacturer (OXOID Ltd., Basingstoke, Hampshire, UK).

All agar plates were prepared in 90-mm Petri dishes with 20 mL of agar, giving a final depth of 4 mm. Overnight broth cultures were prepared in Bacto heart infusion broth which was prepared according to the instructions of the manufacturer (Becton, Dickinson and Co., Sparks, MD), appropriately adjusted in PBS (phosphate-buffered saline) or in Saline 0.9% in order to yield approximately 1.0×10^6 cfu/mL (colony forming units/mL). Whatman paper disks (Whatman International Ltd, Maidstone, UK) of 6-mm diameter were placed on the inoculated agar surfaces and were impregnated with 20 μL of each chemical to be tested.

Standard antibiotics were used in order to provide a control for the sensitivity of the test organisms in the experiments. For each bacterium, two antibiotics (*15*, *16*) were chosen as controls: for *E. coli*, gentamicin and tetracycline, and for both *S. aureus* and *B. subtilis*, gentamicin and vancomycin. Standard graphs for each antibiotic were prepared by testing paper disks containing varying amounts of the antibiotic against a standard organism. The concentrations of each antibiotic used for that purpose were 500, 200, 100, 50, and 10 μg/mL (in 18.2 MΩ water). Each test was performed in duplicate and plates with and without Tween 80 were used for *E. coli*. Standard graphs for each antibiotic and each bacterium were prepared by plotting the logarithm of the concentration of the antibiotic versus the mean zone of inhibition, and a very good linearity was observed. The same procedure was followed for mastic oil, mastic gum, the mastic oil distillation fractions, selected standards (see below), and ethanol. Each sample (20 μL of the liquids and 20 μL of a 30 mg/mL solution of the gum extracts) was applied to the paper disks, and they were tested on plates containing MH agar (without Tween 80).

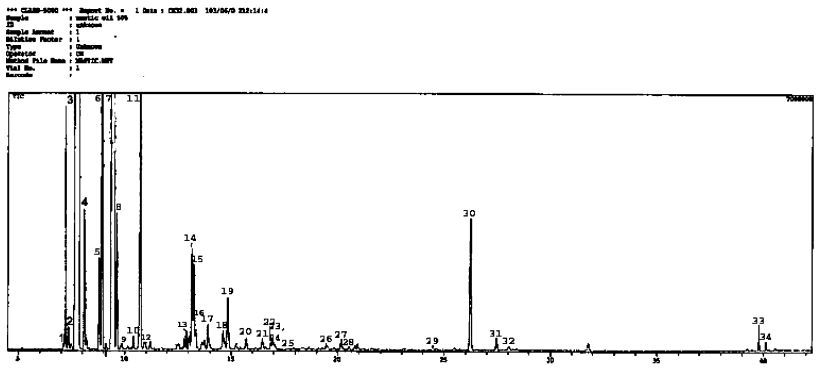
RESULTS AND DISCUSSION

Chemical Composition of Mastic Oil and Gum. The oil, gum, and distillation fractions were analyzed by GC−MS using the method described above. **Table 1** contains the identified peaks of fractions 1−4, along with their percentages, in comparison with those of mastic oil and of the ethanol-soluble components of mastic gum. GC−MS analysis of the oil and the gum led to the identification of the majority of the components, which are listed in Table 1 along with their semiquantitative data. A typical GC−MS chromatogram of mastic oil, obtained with the analytical method described above, is illustrated in **Figure 1**. The GC−FID analysis of mastic oil and gum showed no significant difference compared to that obtained by GC−MS. The identification of the components was based on comparison of their mass spectra with those of NIST12.LIB and NIST62.LIB libraries, as well as on comparison of their retention indices (*17*) and of the standard components analyzed (see Materials and Methods).

Table 1. Chemical Composition of Mastic Oil and Mastic Gum As Determined by GC–MS and GC–FID Analysis and Comparison of Their Chemical Composition with the Collected Fractions, Assuming That the Chromatogram Peaks Areas Are Proportional to the Concentration of Analytes

			percentage (%)[a]					
compound	RI[b]	identification methods[c]	mastic gum	mastic oil	fraction 1	fraction 2	fraction 3	fraction 4
octyl formate	923	MS	–[d]	tr[e]	tr	tr	tr	–
tricyclene	926	MS	–	0.1	–	–	–	–
α–pinene	939	GC–MS	40.9	63.3	66.3	52.6	2.4	3.1
camphene	953	MS	1.0	0.6	1.2	1.4	tr	–
sabinene	976	MS	0.3	0.4	0.8	1.5	0.1	0.2
β-pinene	980	GC–MS	1.7	3.3	5.4	8.8	0.6	1.2
β-myrcene	991	GC–MS	9.0	25.0	13.1	21.6	7.8	17.0
methyl-*o*-cresol	1009	MS	0.3	0.6	1.0	1.8	0.7	1.4
p-cymene	1026	GC–MS	–	0.1	0.1	0.2	0.1	0.3
limonene	1031	GC–MS	0.8	1.5	1.7	3.2	1.8	3.8
(*Z*)-*β*-ocimene	1040	MS	–	tr	tr	0.1	–	–
(*E*)-*β*-ocimene	1050	MS	–	tr	0.1	0.1	tr	0.2
α-terpinolene	1087	MS	–	tr	–	–	–	–
α-pinene epoxide	1095	MS	–	–	1.4	1.4	4.2	7.7
linalool	1098	GC–MS	0.8	0.5	0.4	0.3	2.7	3.2
perillene	1099	MS	0.8	0.5	0.8	0.8	2.6	4.3
cis-verbenol	1100	MS	–	0.1	0.4	0.4	3.5	5.2
α-campholene aldehyde	1126	MS	0.3	0.1	0.1	0.2	0.6	0.9
trans-pinocarveol	1139	MS	–	0.1	0.2	0.2	2.3	2.3
trans-verbenol	1143	MS	0.5	0.3	0.8	0.8	9.6	10.4
β-pinene epoxide	1156	MS	–	–	tr	tr	0.5	0.5
myrtenal	1193	MS	0.5	0.1	0.2	0.2	1.7	1.9
α-terpineol	1180	GC–MS	–	tr	–	–	–	–
myrtenol	1188	MS	–	tr	–	–	–	–
verbenone	1189	GC–MS	0.8	0.1	0.4	0.4	6.5	6.1
trans-carveol	1217	MS	–	tr	tr	tr	0.3	0.1
dihydrocarveol	1226	MS	–	–	–	–	0.2	0.2
neral	1240	MS	–	0.1	–	–	–	–
linalyl acetate	1257	MS	–	tr	tr	tr	0.5	0.2
(*E*)-anethole	1283	GC–MS	tr	0.1	tr	tr	0.5	0.4
α-fenchyl acetate	1290	MS	–	tr	–	–	0.2	0.1
neryl acetate	1365	MS	–	–	0.2	0.2	1.3	1.7
α-copaene	1381	MS	–	tr	–	–	0.2	0.1
β-caryophyllene	1419	GC–MS	5.3	0.9	0.1	0.1	2.5	2.1
α-humulene	1454	MS	0.1	0.1	–	tr	0.4	0.2
(*E*)-methylisoeugenol	1500	GC–MS	0.9	tr	–	–	0.4	0.1
(*Z*,*Z*)-farnesol	1713	MS	11.9	0.1	–	–	6.3	3.2
(*E*,*Z*)-farnesol	1742	MS	0.1	tr	–	–	1.8	1.6

[a] Percentages obtained by GC–FID. [b] Retention index relative to *n*-alkanes on SGE-BPX-5 capillary column (similar type to DB-5 capillary column). [c] Methods: GC, identification based on retention times of standard compounds on SGE-BPX-5 capillary column; MS, tentatively identified based on computer matching of the mass spectra of peaks with NIST12.LIB and NIST62.LIB libraries and published data. [d] An en-dash denotes that the percentage was below the limit of detection. [e] Less than 0.1%.



1. octyl formate
2. tricyclene
3. α – pinene
4. camphene
5. sabinene
6. β – pinene
7. β – myrcene
8. methyl o – cresol
9. (Z)-β-ocimene
10. p-cymene
11. limonene
12. (E)-β-ocimene
13. α-terpinolene
14. linalool
15. perillene
16. cis-verbenol
17. α – campholene aldehyde
18. trans – pinocarveol
19. trans-verbenol
20. neral
21. myrtenal
22. α-terpineol
23. myrtenol
24. verbenone
25. trans – carveol
26. linalyl acetate
27. (E)-anethole
28. α – fenchyl acetate
29. α – copaene
30. β – caryophyllene
31. α – humulene
32. (E)-methyl iso – eugenol
33. (Z,Z)-farnesol
34. (E,Z)-farnesol

Figure 1. GC–MS chromatogram of mastic oil obtained with the method described previously.

The major constituents of the essential oil of *P. lentiscus* var. *chia* were α-pinene (63%), *β*-pinene (3.3%), *β*-myrcene (25%), limonene (1.5%), and *β*-caryophyllene (1%), assuming that TIC (total ion current) as integrated over the peak in the GC−MS chromatogram is proportional to their concentration in the sample. Other constituents therefore account for 6.2% of the total concentration. For the gum, the major ethanol-soluble constituents were the same, but the relative percentages differed from those found in the oil: α-pinene (40%), *β*-pinene (1.5%), *β*-myrcene (9%), limonene (1.0%), and *β*-caryophyllene (5%). This difference is presumably due to the different ways the gum and the oil are produced.

As was intended, the distillation process separated the more volatile components from the less volatile ones. Fraction 1 has a similar chromatogram to that of mastic oil up to where verbenone (**Figure 1**) is eluted, although some peaks are more

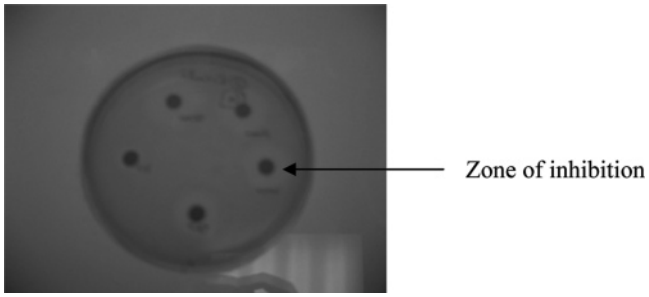


Figure 2. Zones of inhibition of *E. coli* against gentamicin on medium without Tween 80.

enhanced in the first fraction. The later eluting peaks of mastic oil either do not appear in the chromatogram of the fraction or they are very small. Fraction 2 is very similar to fraction 1, but some peaks are enhanced (e.g. β-myrcene, β-caryophyllene) and there is less of the major compound α-pinene. There is a clear enhancement of the later eluting peaks in fraction 3 (which was the first washing of the undistilled fraction) compared to mastic oil and fractions 1 and 2, which is consistent with the aims of distillation procedure. Fraction 4 is very similar to Fraction 3.

Antibacterial Activity of Mastic Oil, Its Components, and Its Fractions. *E. coli* and *S. aureus* were inoculated in agar plates with and without Tween 80, and several interesting observations were made: for *E. coli*, the growth of bacteria was quite similar on MH agar with and without the detergent, which indicated that the latter did not play a significant role. On the other hand, the addition of Tween 80 to the medium inhibited significantly the growth of *S. aureus* for reasons that were not clear. According to Carson et al. (*14*), the reduction in zone size observed when the detergent is added to the MH agar may be due to the Tween 80 allowing better distribution of components through the agar, resulting in a lower overall concentration. Alternatively, the Tween 80 may have enhanced the growth of the test organisms, as it is a source of oleic acid, or it could act as an antagonist to the oil components. Therefore, it was decided that plates containing the detergent would be used mainly in the disk diffusion tests with *E. coli*. As already mentioned, for that bacterium, there was essentially no difference between the detergent and nondetergent plates, suggesting that the solubility of the analytes was not the key issue in the different effectiveness of the analytes.

Figure 2 shows the Zones of Inhibition (ZoI) of growth of *E. coli* against gentamicin on medium without the detergent around the paper disks impregnated with different concentrations of the antibiotic.

The results of the antibacterial assays for *E. coli*, *S. aureus*, and *B. subtilis* are reported in **Table 2**, along with the corresponding amount of antibiotic (gentamicin, vancomycin, and tetracycline) that would give the same zone of inhibition as the chemical. The calculation of the corresponding amount of antibiotic is based on the respective equation of graph when plotting the logarithm of the concentration of the antibiotic versus the mean zone of inhibition.

As seen in **Table 2**, all three bacteria are resistant to α-pinene, which is the most abundant compound of mastic oil (65%), and this is in agreement with literature reference (*21*). A variation in the antibacterial activity of the other tested chemicals against the three bacteria is noticed; *E. coli* is resistant to β-myrcene, while *S. aureus* shows an intermediate response to that chemical and *B. subtilis* is sensitive to its presence. It should be noted that β-myrcene is the compound with the second highest percentage (25%) in the composition of mastic oil. *p*-Cymene, β-caryophyllene, methyl isoeugenol, limonene, γ-terpinene, and *trans*-anethole show only moderate antibacterial activity, and in some cases the bacteria are resistant to them. Furthermore, *E. coli* and *S. aureus* are resistant to β-pinene, while it inhibits only slightly the growth of *B. subtilis*. Verbenone (0.07%), α-terpineol (0.01%), and linalool (0.5%) are some of the trace components of mastic oil, but they show higher antibacterial activity than all other components, which is comparable to that of mastic oil itself.

A number of the fractions collected by microdistillation were also tested for antibacterial properties. Fractions 1 and 3 were chosen, since they were the ones that differed the most from a chemical point of view, and they were tested using the same procedure against the three bacteria. As observed in **Table 3**, fraction 3, which shows an enhancement of the later eluting peaks, has a stronger antibacterial activity than fraction 1, which, on the other hand, shows an enhancement of the earlier eluting peaks. It is, however, interesting that neither of them have the activity that mastic oil shows, although the antibacterial activity of fraction 3 is closer to that of the essential oil compared to the activity of fraction 1. This implies that the compounds that

Table 2. Comparison of Zones of Inhibition of the Three Bacteria against Each Compound and a Corresponding Amount of Antibiotic

	E. coli			*S. aure us*			*B. subtilis*		
compounds	ZoI (mm)	corresponding gentamicin (μg/mL)	corresponding tetracyclin (μg/mL)	ZoI (mm)	corresponding gentamicin (μg/mL)	corresponding vancomycin (μg/mL)	ZoI (mm)	corresponding gentamicin (μg/mL)	corresponding vancomycin (μg/mL)
α-pinene	0	–	–	3.5	–	–	0	–	–
β-myrcene	0	–	–	10	22	84	19.5	97	~4300[a]
p-cymene	0	–	–	3.5	–	–	9	6	34
β-caryophyllene	7	7	6	10.5	27	125	13.5	19	270
verbenone	12	40	42	21.5	~2000[a]	~820000[a]	17	49	~1360[a]
α-terpineol	17	218	292	13	72	9201	15.5	33	680
methylisoeugenol	8	10	9	7	7	8	10.25	8	60
limonene	7	7	6	0	–	–	9.5	7	42
β-pinene	0	–	–	0	–	–	6.75	–	12
linalool	12.5	47	51	14.5	130	~3050[a]	26	556	~87000[a]
γ-terpinene	7	7	6	8	10	17	12	13	135
trans-anethole	7	7	6	8	10	17	10	8	54
mastic oil	12	40	42	18.5	619	~75000[a]	17	49	~1360[a]
mastic gum in EtOH	0	–	–	0	–	–	0	–	–
ethanol	0	–	–	0	–	–	0	–	–

[a] The number is too high and it is off scale of the graph for the antibiotic calibration graph, therefore, is not reliable.

Table 3. Zones of Inhibition of the Three Bacteria against the Fractions of Mastic Oil and against the Essential Oil

	ZoI (mm) of bacteria		
fractions of mastic oil	*E. coli*	*S. aureus*	*B. subtilis*
mastic oil	12	18.5	17
fraction 1	0	7	7
fraction 3	6.5	10	11

contribute the most in the antibacterial activity of mastic oil are the ones that appear in the middle of its chromatogram (e.g. verbenone, linalool, α-terpineol) and are clearly enhanced in fraction 3. This conclusion is in agreement with the observations made from **Table 2** and **Figure 1**. However, we cannot overlook the fact that many of the components of mastic oil contribute to its antibacterial activity.

Reasons for the effectiveness of mastic oil against the bacterium *H. pylori* are of particular interest, given the difficulties encountered in treating such infections. As *H. pylori* and *E. coli* are both Gram-negative rods, the observations made for mastic oil and *E. coli* may be relevant for *H. pylori*.

In conclusion, the antibacterial activity of mastic oil can be attributed to the combination of several components rather than to one particular compound. It is also interesting to note that different bacteria are susceptible or not to different compounds of the essential oil. So it can be assumed that the antibacterial activity of mastic oil against the bacteria tested is due to its cocktail of components, including some of the trace elements, not all of which have been identified in this study.

LITERATURE CITED

(1) Al-Habbal, M. J.; Al-Habbal, Z.; Huwez, F. U. A double-blind controlled clinical trial of mastic and placebo in the treatment of duodenal ulcer. *J. Clin. Exp. Pharmacol. Physiol.* **1984**, *11* (5), 541−4.

(2) Huwez, F. U.; et al. Mastic gum kills *Helicobacter pylori*. *N. Engl. J. Med.* **1998**, *339*, 1946.

(3) Al-Said, M. S.; Ageel, A. M.; Parmar, N. S.; Tariq, M. Evaluation of mastic, a crude drug obtained from *Pistacia lentiscus* for gastric and duodenal anti-ulcer activity. *J. Ethnopharmacol.* **1986**, *15* (3), 271−8.

(4) Iauk, L.; Ragusa, S.; Rapisadra, A.; Franco, S.; Nicolosi, V. M. In vitro antimicrobial activity of *Pistacia lentiscus* L. extracts: Preliminary report. *J. Chemother.* **1996**, *8* (3), 207−9.

(5) Topitsoglou-Themeli, V.; Dagalis, P.; Lambrou, D. A. Chios mastiche chewing gum and oral hygiene. I. The possibility of reducing or preventing microbial plaque formation. *Hell. Stomatol. Chron.* **1984**, *28* (3), 166−70.

(6) Doukas, C. Cosmetics that contain mastic gum and mastic oil. *Chem. Chron.* **2003**, *12*, 36−39.

(7) Magiatis, P.; Melliou, E.; Skaltsounis, A. L.; Chinou, I. B.; Mitaku, S. Chemical composition and Antimicrobial activity of the essential oils of *Pistacia lentiscus* var. *chia*. *Planta Med.* **1999**, *65*, 749−752.

(8) Papageorgiou, V.; Mellidis, A.; Argyriadou, N. The chemical composition of the essential oil of mastic gum. *J. Ess. Oil Res.* **1991**, *3*, 107−110.

(9) Markham, J. Antimicrobial effectiveness of tea tree oil. (http://www.ascc.com.au/private/journal/199908/21.shtml).

(10) Pybus, D. H.; Sell, C. S. The Chemistry of Fragrances. *RSC Paperbacks* 1999.

(11) Finegold, S. M.; Martin, W. J. *Diagnostic Microbiology*, 6th ed.; C. V. Mosby Co.: St. Louis, MO, 1982.

(12) Cruickshank, R. *Medical Microbiology: A Guide to the Laboratory Diagnosis and Control of Infection*, 11th ed.; E. & S. Livingstone Ltd.: London.

(13) Stokes, E. J.; Ridgway, G. L.; Wren, M. W. D. *Clinical Microbiology*, 7th ed.; Edward Arnold: London, 1993.

(14) Carson, C. F.; Riley, T. V. Antimicrobial activity of the major components of the essential oil of *Melaleuca alternifolia*. *J. Appl. Bacteriol.* **1995**, *78* (3), 264−9.

(15) Hammond, S. M.; Lambert, P. A. *Antibiotics and Antimicrobial Action*, 1st ed., Edward Arnold: London, 1978.

(16) BSAC Disc Diffusion Method for Antimicrobial Susceptibility Testing, Version 2, July 2001.

(17) Adams, R. P. *Identification of Essential oil components by Gas Chromatography/Mass Spectroscopy*; Allured Publishing Corp.: Carol Stream, IL, 1995.

(18) Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by agar dilution, European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID), June 2000.

(19) Marone, P.; Bono, L.; Leone, E.; Bona, S.; Carretto, E.; Perversi, L. Bactericidal activity of *Pistacia lentiscus* mastic gum against *Helicobacter pylori*. *J. Chemother.* **2001**, *13* (6), 611−14.

(20) Daferera, D.; Pappas, C.; Tarantilis, P. A.; Polissiou, M. Quantitative analysis of α-pinene and β-myrcene in mastic gum oil using FT−Raman spectroscopy. *Food Chemistry* **2002**, *77* (4), 511−15.

(21) Jeon, H.-J.; Lee, K.-S.; Ahn, Y.-J. Growth-Inhibiting Effects of Constituents of *Pinus* densiflora Leaves on Human Intestinal Bacteria. *Food Sci. Biotechnol.* **2001**, *10* (4), 403−407.

Received for review March 22, 2005. Revised manuscript received July 26, 2005. Accepted July 31, 2005.

JF050639S

Journal of Liquid Chromatography & Related Technologies®, 28: 2995–3003, 2005

ISSN 1082-6076 print/1520-572X online
DOI: 10.1080/10826070500294966

Theoretical Aspects of the Enantiomeric Resolution of Dimetallo Helicates with Different Surface Topologies on Cellulose Columns

Syma Khalid
Department of Biochemistry, University of Oxford, Oxford, U.K.

P. Mark Rodger and Alison Rodger
Department of Chemistry, University of Warwick, Coventry, U.K.

Abstract: Cellulose has been used to separate the enantiomers of a range of dimetallo coordination compounds with different surface topologies. The compounds are all approximately cylindrical in shape, but are based on octahedral coordination at the metals and are also helical. When separation has been achieved, the first eluted enantiomer has always been proven to have a negative circular dichroism (CD) signal for its longest wavelength metal to ligand charge transfer band. In order to understand the underlying basis for the elution order, gas phase molecular dynamics and snap-shot minimisations of each enantiomer with the repeat unit of cellulose, glucose, have been undertaken. For new dimetallo helicates, it is important to have a quick assessment of the enantiomeric identity of the first eluted compound. To this end, the coupled-oscillator model of CD has been applied to relate the signs of the CD signals to the identity of the enantiomers. This correlation is consistent with crystallographic data for the first eluted enantiomer of the parent compound.

Keywords: Dimetallo helicates, Circular dichroism, Cellulose, Chromatography, Molecular modelling, Glucose

Address correspondence to Prof. Alison Rodger, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, U.K.. E-mail: a.rodger@warwick.ac.uk

INTRODUCTION

A series of cationic dimetallo helicates, based on octahedral coordination at the metals by three fairly rigid tetra-dentate ligands, have been designed and synthesised at Warwick.[1–4] These helicates are of appropriate size and shape to target the major groove of DNA,[5] and upon binding they induce unusual structural changes in the DNA.[2,6] The compounds are chiral and the helicity of the complex has been found to have a significant effect on the DNA binding of the enantiomers.[6] After exhaustive searches of established resolution methodologies, we found that the enantiomers of the parent compound $[Fe_2(L_1)_3]^{4+}$ (Figure 1) can be separated using cellulose (in paper or packed in a column) with 0.02 M aqueous sodium chloride as the eluent,[7] and we have subsequently refined the mobile phases to minimise the amount of NaCl required (since this ends up in solution with the resolved enantiomers and is very difficult to remove).[8] The chromatographic behaviour of derivatives of the parent dimetallo helicate differs depending on where substituents are located. The effects, which are given in detail in reference [8], can be summarised with reference to the position labels in Figure 1 as follows. Replacing the bridging CH_2 group with O at position **A** worsens the enantiomeric separation, though does not remove it completely. Adding methyl groups on the imine bond at position **B** destroys the enantiomeric separation. Substitution at positions 3 (**C**), 5 (**E**) and 4 (**D**) of the terminal pyridine results in a small reduction in separation efficiency with their effect being in the order (**C**) < (**E**) < (**D**).

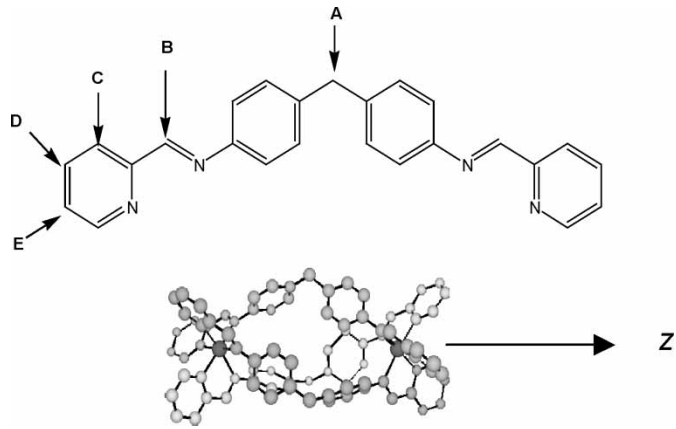


Figure 1. Ligand and dimetallohelicate structures of the parent compound $[Fe_2(L_1)_3]^{4+}$, with sites of potential substitution labelled (the substitution pattern is usually symmetric about A and is assumed to be so in this work). *z* is the 3-fold axis of the helicate.

To date, whenever chiral resolution has been achieved, the first eluting compound gives a negative circular dichroism (CD) signal in its longest wavelength band. The first eluted band of the parent compound has been shown by X-ray crystallography to be the M-helicate.[8] In order to be able to use the methodology and enantiomer assignments suggested by the circular dichroism, it is necessary to understand the basis of the molecular discrimination on the column and also the mechanistic origin of the CD signal. The aim of the work reported in this paper was, therefore, to understand why (or why not) we observed a chiral discrimination when the dimetallo helicates were eluted from a cellulose (polymer of glucose) stationary phase by examining the interactions between parent helicate and a glucose monomer, and to understand the basis of the CD spectra which could be obtained for the two enantiomers.

EXPERIMENTAL

Spectroscopy

$[Fe_2(L_1)_3]Cl_4$ was synthesised and resolved as described in reference.[6] The CD spectra were collected in 1 cm pathlength cuvettes using a Jasco J-715 spectropolarimeter, and averaged over 4 scans with response time = 1 s. Concentrations were determined using $\varepsilon_{574\,nm} = 16{,}900\,mol^{-1}\ dm^3\ cm^{-1}$ for the metal-ligand charge transfer band.[2,8] The absorption spectra were collected on a Jasco V-550 spectrometer using a 1 cm pathlength cuvette.

Computational Modelling

Models of glucose and the two helicate enantiomers were constructed. High temperature molecular dynamics (MD) simulations were performed with the cylinder frozen, while the glucose was mobile. These simulations were designed to identify likely sugar interaction sites on the cylinder. It was hoped they would provide molecular level insights into the differential cellulose binding of the two enantiomers. The configurations with the most favourable energies from the MD simulations were then energy-minimised to further optimise the geometries. The helicate coordinates were taken from the crystal structure, while glucose was constructed using the 2D sketcher module in Quanta and then energy-minimised. Compounds were modelled using the CHARMM22 all-atom force-field with modifications, so the glucose remained in the chair conformation.[10–12] Simulations were performed using DL_POLY. The conversion from CHARMM to DL_POLY force-field formats was achieved using a purpose-built program that interpreted the CHARMM prm and crd files. Simulations were performed in the *NVT* ensemble at 900 K using the Hoover thermostat with a time constant

of 0.1 ps and a time step of 2 fs. The high temperature was adopted to ensure the glucose moved around sampling all potential interaction sites.

RESULTS

Circular Dichroism Spectroscopy

Spectra for the two 'baseline separated' bands from a racemic mixture of $[Fe_2(L_1)_3]^{4+}$, loaded on a column packed with cellulose and eluted with 20 mM NaCl[6], are given in Figure 2. CD spectra for all compounds where the enantiomers could be resolved[8] give the same sign pattern, from long wavelength to shorter wavelength of $-/+/-/+/-/+$ for the first eluting enantiomer. The second eluting enantiomer has the opposite sign pattern. The labelling of the spectra in Figure 2 with M and P follows from the CD theory outlined below. This labelling has recently been confirmed for the first eluting enantiomer of the parent compound only, from a crystal structure of a crystal grown from column eluent.[8] The CD theory given below enables us to generalise the crystallographic assignment of the parent compound to all derivative compounds.

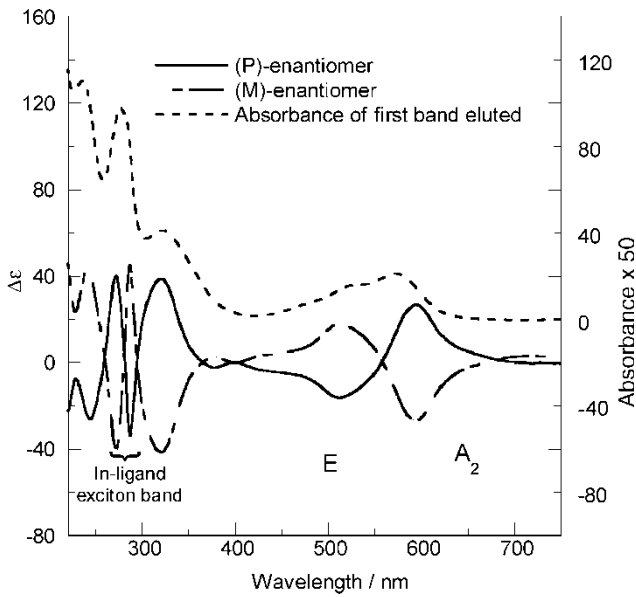


Figure 2. Absorbance of the first eluting $[Fe_2(L_1)_3]^{4+}$ band (mobile phase: 0.02 M NaCl) and the CD spectra of both bands (converted to $\Delta\varepsilon$ using concentrations determined from the absorbance magnitude and the extinction coefficient).[2,8] Absorbance (times 50) scale is on the right hand side. The spectrum labelled M is the enantiomer eluted first. Labels M and P are assigned on the basis of the CD analysis of this work.

In order to be able to use CD spectroscopy as a fast means of confirming the handedness of fractions eluted from the cellulose columns for all the dimetallo helicates, it is important to understand the origin of the signs of the bands and to be able to relate this to the handedness of the helicate. The relevant theoretical model for CD intensity of electric dipole allowed transitions in independent chromophores is the exciton coupling model.[15] For degenerate transitions, if this model is appropriate, one expects to see a characteristic exciton spectrum with a very sharp transition from negative to positive (or vice versa) signals. Such a signal is apparent in the dimetallo helicate spectra of Figure 2 centered at 280 nm. This region of the spectrum is dominated by the in-ligand transitions on the three identical chelates.

As the conjugation of the π-chelate of the dimetallo-helicates is broken at the central bridge, we can spectroscopically model the helicates as two tris-chelate metal complexes stacked vertically. The CD arising from the in-ligand couplings of single-metal tris-chelate complexes is well established.[14–16] The equations for the bimetallo tris-chelate CD can therefore be taken from the literature[15,17] after multiplying by a factor of two. As tris-chelates have $\mathbf{D}_3$ symmetry, their transitions are either polarised along the 3-fold axis, denoted z in Figure 1, or in the plane perpendicular to this. The in-ligand transitions of relevance for CD are all polarised along the chelate backbone, as the ones perpendicular to this do not give rise to a CD signal (the electron movement is in a plane, so not chiral). The z-polarised transitions are by convention labelled A_2 and the x/y-polarised, E. The CD intensity, R, for the A_2 and E polarised in-ligand bands of the P bi-metallo helicate is, therefore,

$$R\left(\frac{z}{x/y}\right) = R\left(\frac{A_2}{E}\right) = \pm\frac{\varepsilon\mu^2\rho}{\sqrt{2}\hbar} \tag{1}$$

(where the upper, in this case positive, sign refers to the upper transition polarization, namely z) with transition energies

$$\varepsilon\left(\frac{z}{x/y}\right) = \varepsilon\left(\frac{A_2}{E}\right) = \varepsilon \pm \frac{\mu^2}{12\sqrt{3}\rho^3} \tag{2}$$

where ε is the energy of the unperturbed transition, μ is the magnitude of its transition dipole moment, and ρ is the distance from the metal to the centre of a chelating part of the ligand. As $R(z)$ is the z (or A_2) polarized CD that results from the in-phase coupling of the transition moments on the three ligands, the A_2 band occurs at the higher energy (shorter wavelength). It is positive in sign for the P-enantiomer. $R(x/y)$ by way of contrast results from the out-of-phase coupling of the transition moments and is opposite in sign from $R(z)$. The energy ordering is the same for the M-enantiomer, but the sign pattern is inverted. This lets us identify the handedness of enantiomers eluting from the column: if the higher energy (shorter wavelength) band of the 280 nm

couplet is positive, then the complex is the P enantiomer. Thus, the second eluting enantiomer is P.

The other bands in the spectrum are less clearly excitonic in nature, arising as they do from a complicated overlay of different metal ligand charge transfer (MLCT) and in-ligand bands. However, it is convenient to have an empirical rule based only on the longest wavelength CD band. To this end, it is necessary to understand the longest wavelength MLCT band CD. The dilemma here is that the available data on transition polarisations of long wavelength charge transfer transitions for low spin iron(II) and ruthenium(II) (including our own stretched film *LD* assignments of transition polarisations in Fe_2L_3)[16] show significantly more (>90%) E than A_2 intensity of the long wavelength end of the absorbance spectrum, but the CD spectra show similar magnitude positive and negative bands, which, assuming the CD signals arises from a similar mechanism, requires both A_2 and E absorption intensities to be of similar magnitudes.

The dilemma is resolved by realising that E-polarised MLCT transitions occurring from an Fe metal *d* orbital to the middle of a ligand π^* orbital[18] will have significant absorbance intensity, but will involve planar electron movement, and thus will not give rise to a CD signal MLCT transitions into long axis polarised π^* states will give rise to both E and A_2 MLCT absorbance intensity, both of which will have a helical twist and give rise to a CD signal. The net effect will be that the absorbance spectrum is dominated by E polarised transitions in regions where there are short-axis polarised MLCT bands, whereas the CD spectrum will have similar magnitude contributions from the two polarisations in accordance with the experimental observation. Using the enantiomeric assignment made using the 280 nm in-ligand band (for which we also know the energy ordering of components, see above), we conclude that the longest wavelength P enantiomer MLCT CD, which is positive in sign, is due to an A_2 transition (Equation (1)). The longest wavelength band for all helicates is thus A_2, and can be used to assign the handedness of the helicates. For all the compounds investigated to date,[8] if enantiomeric resolution occurs, the first eluting enantiomer has a negative CD sign, M.

Modelling

The molecular modelling of glucose, the cellulose monomer, interacting with the two enantiomers of $[Fe_2(L_1)_3]^{4+}$ revealed three sites with more favourable interaction energies. Two of these involved the sugar being near one of the metal centres and the other was near the bridging CH_2 group (Figure 3). The interaction energies of each enantiomer of $[Fe_2(L_1)_3]^{4+}$ with a sugar in each of these sites are given in Table 1. Interaction energy = total configurational energy − configurational energy of glucose (the helicate is frozen so only contributes to the non-bonded energy). Simulations were also performed for the helicate where the CH_2 bridge is replaced by O.

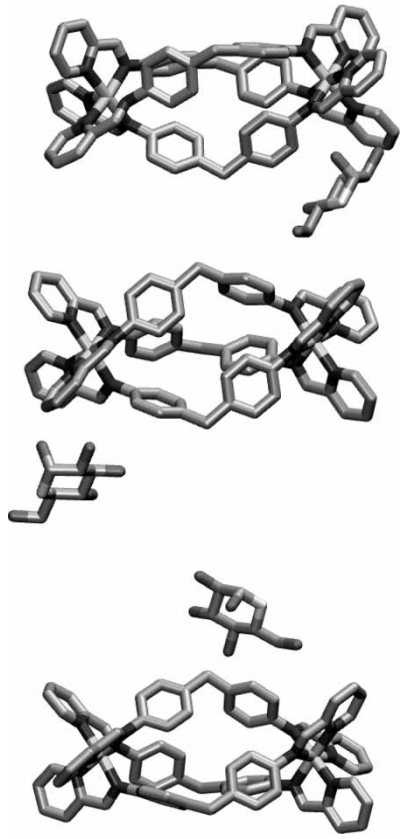


Figure 3. Site 1 (top), site 2 (middle) and site 3 (bottom). Hydrogens have been omitted for clarity. The sites are shown for the M enantiomer.

In each of the three interactions sites on $[Fe_2(L_1)_3]^{4+}$ favoured by the glucose, the interaction is strongest between the P-enantiomer and the glucose, meaning that the M-enantiomer is expected to come off a column first—as it indeed does. It follows that, when sites 1 and 2 are blocked by, e.g., a methyl substituent on the immine bond, the resolution between the enantiomers is expected to be reduced. In practice,[8] loss of sites 1 and 2 is sufficient to completely remove the enantiomeric resolution. The calculations suggest, however, that with great care and collecting beginnings and ends of fractions, some degree of resolution might be achieved.

The overall strongest interaction for the glucose with the P-enantiomer of $[Fe_2(L_1)_3]^{4+}$ is at the site nearest the bridging group (site 3, Fig. 3). Contrary to our expectations (given the charged nature of the metallo-helicate), when the bridging CH_2 was replaced by the potentially H-bonding O atom the interaction of the glucose with site 3 became less attractive, and it always

Table 1. Interaction energies of each enantiomer of $[Fe_2(L_1)_3]^{4+}$ and the analogous compound where the CH_2 bridge is replaced by O (denoted O in the table) with a glucose molecule in each of the 3 favourable sites. Site 3 is near the bridging CH_2 or O

	Site 1	Interaction energy (kcal mol^{-1})	Site 2	Interaction energy (kcal mol^{-1})	Site 3	Interaction energy (kcal mol^{-1})
CH_2	P	−9.0	P	−8.4	P	−11.0
CH_2	M	−7.9	M	−7.1	M	−6.2
O	P	−8.9	P	−8.4	P	—
O	M	−7.9	M	−7.1	M	—

migrated to sites 1 or 2. The experimental consequence of this would be predicted to be that the discrimination between the enantiomers is reduced and their residence time on the column is shortened. This is indeed the case. The enantiomeric resolution of the O-bridged helicates is significantly smaller than that of the parent compound,[8] and both O-bridged enantiomers have larger R_f values (distance moved by analyte divided by distance moved by solvent front) on cellulose paper than the corresponding CH_2 compound.[8]

CONCLUSIONS

From the analysis of the CD of the in-ligand exciton band and the longest wavelength MLCT region of the spectrum, we can conclude that CD can be used to assign the handedness of the di-iron triple helicates. A negative long wavelength band corresponds to the M-enantiomer. To date, from the data in reference [8], we conclude here that it is always the M-enantiomer which elutes first from cellulose and the P-enantiomer which elutes second, whenever any resolution is achieved. For other metals initial consideration of the in-ligand region of the spectrum would be necessary before establishing such a rule. Molecular modelling of glucose monomers and both M and P helicates has been used to understand why the elution occurs in this way. In general terms, P interacts more favourably with glucose so is retained on a column. Much of this discrimination arises from interaction near the metal atoms. It follows from this, that when this site is filled up with a substituent on the immine, such as CH_3, then the chiral resolution is reduced or removed. The origin of lower resolution when the bridging atom is changed from CH_2 to the electronegative O arises, somewhat surprisingly, from the fact that the H-bonding glucose does not interact as favourably with the O-bridged site as with the other sites near the metal. This then reduces the time the helicates spend on the column and reduces the discrimination between the enantiomers, as there is no contribution from the bridging site.

REFERENCES

1. Hannon, M.J.; Painting, C.L.; Jackson, A.; Hamblin, J.; Errington, W. Chem. Commun. **1997**, 1807–1808.
2. Rodger, A.; Sanders, K.J.; Hannon, M.J.; Meistermann, I.; Parkinson, A.; Vidler, D.S.; Haworth, I.S. Chirality **2000**, *12*, 221–236.
3. Hannon, M.J.; Painting, C.L.; Alcock, N.W. Chem. Commun. **1999**, 2023–2024.
4. Uerpmann, C.; Malina, J.; Pascu, M.; Clarkson, G.J.; Moreno, V.; Rodger, A.; Grandas, A.; Hannon, M.J. Chem. Eur. J. **2005**.
5. Hannon, M.J.; Moreno, V.; Prieto, M.J.; Molderheim, E.; Sletten, E.; Meistermann, I.; Isaac, C.J.; Sanders, K.J.; Rodger, A. Angew. Chem. Int. Ed. **2001**, *40* (5), 880–884.
6. Meistermann, I.; Moreno, V.; Prieto, M.J.; Molderheim, E.; Sletten, E.; Khalid, S.; Rodger, P.M.; Peberdy, J.C.; Isaac, C.J.; Rodger, A.; Hannon, M.J. Proc. Natl. Acad. Sci. **2002**, *99* (8), 5069–5074.
7. Hannon, M.J.; Meistermann, I.; Isaac, C.J.; Blomme, C.; Aldrich-Wright, J.R.; Rodger, A. Chem. Commun. **2001**, 1078–1079.
8. Peberdy, J.C.; Reudegger, V.; Kerchoffs, J.; Rodger, A.; Hannon, M.J. Submitted.
9. Accelrys. In *QUANTA, Accelrys*; San Diego, CA, 2000.
10. MacKerell, A.D., Jr.; Bashford, D.; Bellot, M.; Dunbrack, R.L., Jr.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnik, S.; Ngo, T.; Nguyen, D.T.; Prodhom, B.; Reiher, W.E., III; Roux, B.; Schlenkrich, M.; Smith, J.C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J. Phys. Chem. B. **1998**, *102*, 3586–3616.
11. Astley, T.; Birch, G.G.; Drew, M.G.B.; Rodger, P.M. J. Phys. Chem. A **1999**, *103*, 5080–5090.
12. Astley, T.; Birch, G.G.; Drew, M.G.B.; Rodger, P.M.; Wilden, G.R.H. Food Chemistry **1996**, *56*, 231–240.
13. Smith, W.; Forester, T.R.J. Mol. Graphics **1996**, *14*, 136–141.
14. Bosnich, B. Acc. Chem. Res. **1969**, *2*, 266–273.
15. Rodger, A.; Nordén, B. *Circular Dichroism and Linear Dichroism*; Oxford University Press: Oxford, 1997.
16. Holder, E.; Trapp, G.; Grimm, G.; Schurig, V.; Lindner, E. Tet. Asym. **2002**, *13*, 2673–2678.
17. Armstrong, D.W.; DeMond, W.; Czech, B.P. Anal. Chem. **1985**, *57*, 481–484.
18. Green, J.M.; Jones, R.; Harrison, R.D.; Edwards, D.S.; Glajch, J.L. J. Chromatogr. **1993**, *635*, 203–209.

Received June 1, 2005
Accepted June 30, 2005
Manuscript 6667

Validation of new microvolume Couette flow linear dichroism cells

Rachel Marrington,[a] Timothy R. Dafforn,[b] David J. Halsall,[c] James I. MacDonald,[d] Matthew Hicks[a] and Alison Rodger*[a]

Received 5th May 2005, Accepted 28th September 2005
First published as an Advance Article on the web 14th October 2005
DOI: 10.1039/b506149k

Long molecules such as fibrous proteins are particularly difficult to characterise structurally. We have recently designed a microvolume Couette flow linear dichroism (*LD*) cell whose sample volume is only 20–40 μL in contrast to previous cells where the volume of sample required has typically been of the order of 1000–2000 μL. This brings the sample requirements of *LD* to a level where it can be used for biological samples. Since *LD* is the difference in absorption of light polarised parallel to an orientation direction and perpendicular to that direction, it is the ideal technique for determining relative orientations of subunits of *e.g.* fibrous proteins, DNA–drug systems, *etc.* For solution phase samples, Couette flow orientation, whereby the sample is sandwiched between two cylinders, one of which rotates, has proved to be the optimal technique for *LD* experiments in many laboratories. Our capillary microvolume *LD* cell has been designed using extruded quartz rods and capillaries and focusing and collecting lenses. We have developed applications with PCR products, fibrous proteins, liposome-bound membrane proteins, as well as DNA–dye systems. Despite this range of applications, to date there is nothing reported in the literature to enable one to validate the performance of Couette flow *LD* cells. In this paper we establish validation criteria and show that the data from the microvolume cells are reproducible, vary by less than 1% with sample reloading, follow the Beer–Lambert law, and have signals linear in voltage over a wide voltage range. The microvolume cell data are consistent with those from the large-volume cells for DNA samples. Surprisingly, upon extending the wavelength range by adding the intercalator ethidium bromide, the spectra in the microvolume and large-volume cells differ by a wavelength dependent orientation parameter. This wavelength variation was concluded to be the result of Taylor-vortices in the large-volume cells which have inner rotating cylinders in our laboratory. Thus the microvolume *LD* cells can be concluded to provide better data than our large-volume *LD* cells, though the latter are still to be preferred for titration series as it is extremely difficult to add sample to the capillary cells without introducing artefacts.

Introduction

Linear dichroism (*LD*) is a differential polarised absorption spectroscopy technique in which molecules in a sample are oriented and then the difference in absorption of light polarised in orthogonal directions (one of which is the orientation direction) is determined. Thus it is the ideal technique for studying classes of long biomacromolecules for which it is intrinsically difficult to gain structural information using more commonly available structural characterisation techniques such as NMR and X-ray diffraction. Recently we have reduced the sample requirements for *LD* by 2 orders of magnitude which has opened up a range of new applications for important biomacromolecules including PCR products,[1] fibrous proteins[2,3] (following early work in the 1970's (*e.g.* ref. 4, 5) but extending the wavelength and therefore the information gained), liposome-bound membrane proteins[6] as well as the more commonly studied DNA-dye systems *e.g.* ref. 7–11. Different methods of *LD* sample presentation exist for different types of sample, for example stretched film for small molecules, flow orientation, squeezed gel and electric field orientation for larger molecules.[6,9,12,13] However, for solution phase samples Couette flow orientation is most commonly used for the study of long biological macromolecules.

Couette flow *LD*, despite being used for structural studies for over 30 years, has not seen wide application, mainly because the required sample volume until recently made it very expensive to undertake experiments with biomacromolecules other than readily available DNAs. In addition few laboratories have been able to build their own cells. However, with our recent advances in Couette cell design both these issues have been solved.[14] Thus it is appropriate to consider the issue of *LD* cell validation to establish protocols before wide spread use in the biochemistry community is established. That is the aim of this paper.

Couette cells (Fig. 1) derive from the work of Maurice Frédéric Alfred Couette and Henry Reginald Arnulph Mallock, who in the late 1800's independently developed a

[a]*Department of Chemistry, University of Warwick, Coventry, UK CV4 7AL. E-mail: a.rodger@warwick.ac.uk; Fax: +44-24-76575795; Tel: +44-12-76574696*
[b]*Biosciences, University of Birmingham, Edgbaston, Birmingham, UK B15 2TT*
[c]*Department of Clinical Biochemistry, Box 232, Laboratory block Level 4, Addenbrooke's Hospital, Hills Road, Cambridge, UK CB2 2QQ*
[d]*Department of Mathematics, University of Warwick, Coventry, UK CV4 7AL*

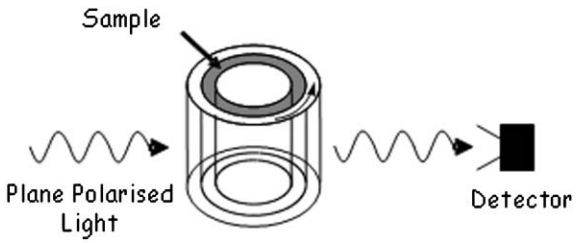


Fig. 1 Schematic diagram of *LD* Couette cell.

means to measure viscosity based on shearing a liquid between coaxial cylinders.[15–18] Since then there have been many applications of the Couette principle within the engineering and physics communities, the most notable example being that of Taylor.[19–21] Couette flow was developed further by Wada and Kozawa in 1964 into an apparatus for UV dichroism measurements.[22] In the 1960's–70's, flow *LD* was recognised as a useful technique and was used in the characterisation of fibrous proteins such as actin as well as probing the orientation of DNA and bound ligands.[23–28] The limitations of instrument design and technology with respect to the required sample volume (1–2 mL) and spectrometer design made it difficult to develop the technique further. Therefore Couette flow *LD* has been mainly restricted in application to DNA and DNA–drug systems as noted above.

The annular gap in Couette *LD* cells has typically been ~500 μm, though experiments using a 50 μm annular gap have been reported[2,5,6] which reduced the sample volume to 200 μL. The sample volumes required are still too large for applications such as analysis of PCR reactions, analysis of expensive biological samples including membrane proteins in liposomes, fibrous peptides, and proteins. This was the motivation for the new cell design which has also been used to develop the new technique of fluorescence detected flow *LD*[13] and to study the bacterial fibrous protein FtsZ.[29] In this paper we report experiments undertaken to show that the data from the new design of *LD* cell which requires only 25 μL sample are consistent with or better than those achieved with more traditional *LD* cells. The limitations of the new cells are also established.

Materials and methods

Design and construction of microvolume Couette *LD* cell

Two microvolume Couette *LD* cells have been designed and developed based on the same design, one being room temperature use only and the other having a heated metal jacket surrounding the capillary housing to enable steady-state thermostatting. The key features of the design are the use of an extruded quartz capillary (~5 mm outer diameter (o.d.) and ~3 mm inner diameter (i.d.)) as the outer rotating cylinder (Fig. 1), and a centrally mounted extruded quartz rod (o.d. dependent on annular gap and hence path length required, typically ~2.5 mm) as the inner stationary cylinder. The capillaries and rods were supplied by Enterprise-Q, Manchester, UK. The microvolume Couette *LD* cell was designed and built in collaboration with Crystal Precision Optics, Rugby, UK. The cell unit has been mounted on a base plate that has been designed specifically for Jasco *CD* spectropolarimeters with large sample compartments, though modifications to small compartments and other instrument bases are straightforward.

The main body of the microvolume Couette *LD* cell has been manufactured from laboratory and food industry specification stainless steel (Fig. 2). The bearings and drive spindle are also made from stainless steel and are designed to be dust and water resistant. The capillary units are demountable for removal during cleaning and sample loading. In this work the capillaries are sealed at the base with Araldite Rapid® and held in position in the metal base unit by an 'O'-ring. A motor is used to drive the 'O'-ring and in turn the capillary. This is controlled electronically by an EP-603 (0–30 V) power supply, adapted to allow more precise measurements (two decimal places) of applied voltage by the addition of a 10-turn potentiometer. A quartz rod is held within a Teflon® unit that is inserted into the metal lid. Before operation, the complete lid unit is inserted into the capillary and screwed securely into place (Fig. 2). This design enables different diameter rods to be used within the microvolume Couette cell.

Fig. 3(a) shows a photograph of the microvolume Couette flow *LD* cell and Fig. 3(b) shows a schematic diagram of the

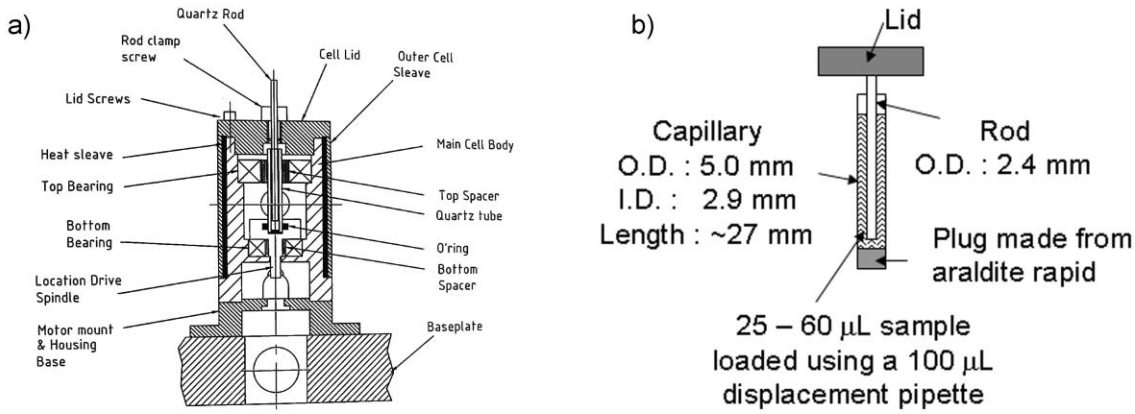


Fig. 2 (a) General arrangement drawing showing detail of quartz microvolume Couette flow *LD* cell and (b) schematic diagram showing the capillary and rod assembly in the microvolume Couette *LD* cell.

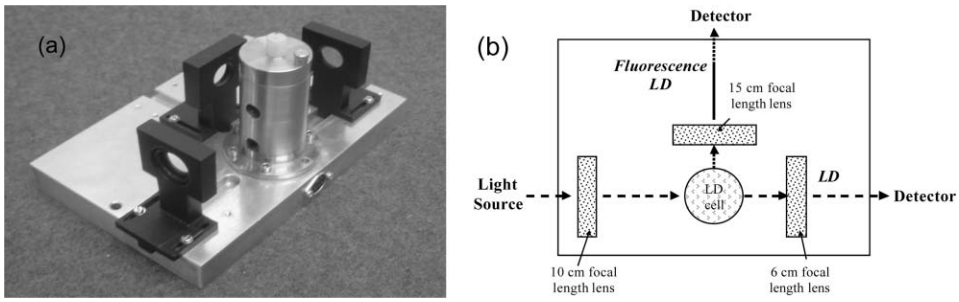


Fig. 3 (a) Photograph of microvolume Couette flow *LD* cell and (b) schematic diagram of microvolume Couette flow *LD* cell. Normal use has the detector on the right hand side; fluorescence-detected *LD* has the photomultiplier tube relocated to the top position.

unthermostatted cell (including the option for fluorescence detection[14]). A 2.54 cm diameter × 10 cm focal length uncoated lens (supplier Edmund Optics®) is placed 10 cm in front of the centre of the rod to ensure that the light is only incident on the middle of the capillary; and a second post-sample focusing lens (2.54 cm diameter × 6 cm focal length) is placed after the sample to focus the diverging light beam onto the photomultiplier tube (this has the added advantage of reducing the loss of light due to scattering by the large molecules often present in the samples for *LD* experiments). The base unit of the microvolume Couette *LD* cell has three slideways to hold the lens holders and to enable them to be finely adjusted by linear movements.

Thermostatted microvolume Couette LD cell

The thermostatted microvolume Couette *LD* cell is a development of the microvolume Couette *LD* cell, with the main difference being the addition of a thermostatted metal jacket with silicon rubber and fibreglass heater pads wrapped around the capillary unit, a temperature probe and a controller. The temperature controller was calibrated to allow the temperature to be set within the operating range (room temperature to ~60 °C). Once a temperature has been programmed, the controller will allow input current to the heater pad to raise the temperature up to near this temperature. The temperature probe is located on the top surface of the capillary unit, and measures the temperature of the steel housing (note, this is not the centre of the cell). When the preset temperature is close to being achieved, the controller switches from a continuous current input to an intermittent current input which holds the set temperature to ±0.2% of the set temperature. Higher temperatures require a longer equilibration time. For example in order to heat the capillary to 63 °C, the cell and temperature unit had to be left for an hour and a half (Fig. 4). The motor should be turned on during the equilibration period as the heat generated from the motor causes a net temperature increase in the sample of ~3 °C, and this should be taken into account when setting the temperature. For example if a temperature of 37 °C is required the temperature controller should be set to 34 °C.

Large-volume Couette LD cell

The large-volume 500 μm annular gap cells used for comparison purposes included the original one (LV Cell 1) from our laboratory, whose design is reported in ref. 12 and a more recently designed version of this (LV Cell 2) where the outer stationary cylinder is a quartz capillary held in place within a stainless steel housing, with gaps for the light to pass through the sample, and a central rotating quartz cylinder.

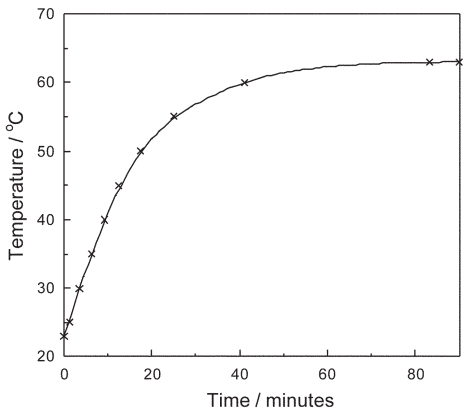


Fig. 4 The temperature of the thermocouple as a function of time from switching on the heater in the thermostatted microvolume Couette *LD* cell.

Experiments for validation of performance of Couette flow LD cells

Calf thymus-DNA (ct-DNA highly lyophilised from Sigma) was dissolved in high purity water (18.2 mΩ) and used for most experiments summarised below. For investigating the larger wavelength range, 750–200 nm, DNA–ethidium bromide solutions were used at concentrations specified below. For LV Cell 1 2000 μL and for LV Cell 2 1000 μL of sample were used at a voltage of 2 V (unless otherwise stated). For the microvolume Couette *LD* cell, samples were individually prepared with volumes of 40 μL placed into the capillary and a voltage of 4 V was used in all experiments, unless otherwise stated. A Gilson air displacement pipette (P100) was used for the loading of capillaries. Volumes of 25–60 μL can be used in conjunction with a ~2.5 mm outer diameter rod. A Jasco J-715 with large sample compartment that has been adapted for *LD* measurements was used for all *LD* experiments. Instrumental parameters for experiments with ct-DNA alone and ct-DNA–ethidium bromide, in both full wavelength (spectrum measurement) and kinetics (time course measurement) are given in Table 1.

Table 1 Parameters used on the Jasco J-715 spectropolarimeter for both wavelength scanning measurements and time course measurement. Parameters in parentheses were those used for calculation of revolutions per minute (rpm) for the microvolume Couette *LD* cell

	DNA alone	DNA–ethidium bromide
Spectrum measurement		
Wavelength/nm	350–190	600–190
Scanning speed/nm min^{-1}	200	500
Response/s	0.5	0.25
Data pitch/nm	0.5	0.5
Band width/nm	2.0	2.0
No. of accumulations	4	4
Time course measurement		
Time/s	0–900 (0–60)	
Response/s	0.5 (0.5 ms)	
Data pitch/s	1 (10 ms)	
Band width/nm	2.0	
Wavelength monitored/nm	259	
No. of accumulations	1	

Error analysis for both the large-volume Couette *LD* cell and microvolume Couette *LD* cell were carried out using both the time course measurement and spectrum measurement programs within the Jasco software. The error in a single analysis was determined by calculating the relative standard deviation (RSD) across all data points within a time envelope of 15 minutes. Multiple analyses were simulated in the LV Couette *LD* cells by stopping and starting the power supply, and in the microvolume Couette *LD* cell by stopping the power supply, emptying the capillary, and then refilling it with the same DNA solution. This experiment was repeated 15 times using full-wavelength analysis and the RSD of the signal maximum at 259 nm (LD_{259}) was calculated. The effect of bandwidth was ascertained by visual inspection of full-wavelength spectra of ct-DNA with different bandwidths and all other parameters as in Table 1. Rotation speeds were calculated for the microvolume Couette *LD* cell by filling the capillary with ct-DNA (200 μM) and then marking the outside of the capillary with a marker pen. The frequency with which the light was interrupted during a time course measurement program (using the parameters in parentheses in Table 1) was then determined. (The mark passes through the light beam twice per revolution.) The supplied voltage was increased in 0.5 V increments from 0 V to 6 V. Different concentrations of ct-DNA were prepared for evaluation of the Beer–Lambert law, ranging from 0–1000 μM.

Concentrations of ct-DNA were determined spectrophotometrically using the molar absorption coefficient per base of $\varepsilon_{259} = 6600\ M^{-1}\ cm^{-1}$.[30] This in turn, along with absorbance data measured independently and scaled for path length differences, was used to determine the sample orientation parameter (S), by rearrangement of eqn (1) where LD^r is the reduced linear dichroism and A the absorbance. The angle between the macroscopic orientation axis and the transition moment termed α was assumed to be 86°.[7]

$$LD^r = \frac{LD}{A} = \frac{3}{2}S\left(3\cos^2\alpha - 1\right) \quad (1)$$

Absorbance spectra in a 1 cm path length quartz cuvette were also collected for all solutions to enable LD^r values to be calculated. A Jasco V-550 UV spectrometer was used for absorbance measurements using the same wavelength and data pitch as that for *LD*.

Results and discussion

Baselines

The capillaries were found to be optically uniform by measuring baseline spectra with different capillary orientations. Thus it is possible to collect a baseline spectrum either by simply stopping the rotating capillary (and hence the alignment force) or using a water–buffer solution with the cell rotating. This situation is in contrast to LV Cell 1 and LV Cell 2 where the baseline measurement has to average over the rotating cell due to variation in the quartz cores as they rotate. It should be noted that it is important to place the rod in the capillary in the same orientation on each occasion as the rod is not as optically uniform as the capillary. If one is working with a capillary cell with light scattering samples, measuring the baseline on the non-rotating sample usually produces better spectra.

Sample volume

Fig. 5 shows that the same *LD* signal has been achieved in the capillary cell with all sample volumes from 25–60 μL. However, it should be noted that the useable volume range is entirely dependent on the amount of Araldite Rapid® used and the location of the rod. The required minimum volume can be calculated for a given capillary from the height of the light beam, the internal volume of the capillary to that height, and the volume of the rod to that height (using capillary internal diameters and rod outer diameters as measured using a micrometer). However, a simple experiment of a number of independently loaded samples should be used to check this and the reproducibility of the loading procedure. Volumes of ~40 μL are used for most experiments reported herein as the smaller volumes are more prone to being loaded with bubbles. Because the capillaries are removable, it is possible to recover the sample.

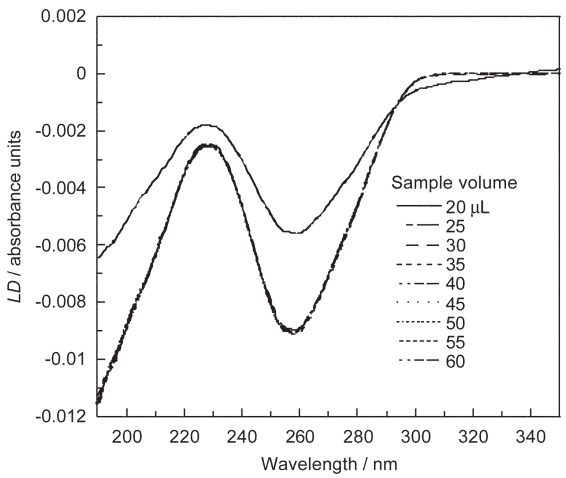


Fig. 5 *LD* spectra of different volumes of ct-DNA (~200 μM) in a capillary.

Sample loading

A series of experiments with ct-DNA showed that not only is a flow gradient established upon rotation of the capillary, but a concentration/size gradient of matter is created. We observed that when the rod is removed, the sample it lifts out is more concentrated than what is left behind. Therefore, all samples should be individually prepared with great care for analysis or reloading. For a series of experiments it is advisable to use the same capillary and rod.

Voltage and rpm relationship

The relationship between voltage and rpm was established by counting the number of revolutions of the capillary by monitoring the period in the HT voltage trace (which indicates absorbance) of a marked capillary, taking into account that the marker will pass through the light beam twice in one revolution (Fig. 6). A linear relationship was observed between rpm and voltage at voltages above 3 V (which fortuitously corresponds to 3000 rpm). It was also found that the signal to noise ratio improves with increasing voltage, with the optimum being 4 V or greater if the sample allows (*i.e.* no bubble formation or turbulence, see below). The minimum voltage of 3 V for a linear relationship between rpm and voltage will depend on individual motors, units and systems under investigation. 4 V is used in this work unless otherwise stated.

Bandwidth dependence of spectra

Bandwidths of 0.1; 0.2; 0.5; 1.0; 2.0; 5.0 and 10.0 nm with the Jasco J-715 were investigated using ct-DNA and the capillary *LD* cell. All bandwidths gave the same spectral shape but it was observed that only 1.0; 2.0 and 5.0 nm bandwidths gave spectra of the same intensity (within tolerances of 1%). There was a large decrease in spectral intensity when using a 10 nm bandwidth (presumably because the size of the light beam is wider than the rod in the far UV region), and the signal to noise ratio was worse when using smaller bandwidths (0.1–0.5 nm). A bandwidth of 2 nm is recommended for use with the microvolume Couette *LD* cell and Jasco spectropolarimeters.

Dependence of capillary LD signal on sample concentration

The validity of the Beer–Lambert law was investigated by measuring LD_{259} of different concentrations of ct-DNA (Fig. 7) in the capillary *LD* cell. It can be seen that all the data points lie on a straight line which indicates that the Beer–Lambert law is obeyed for *LD* of ct-DNA with concentrations up to 1000 μM at 4 V.

Full wavelength range data

Ethidium bromide bound to ct-DNA was used to show that *LD* spectra could be collected in the visible region of the spectrum. A titration series (samples individually prepared) of ethidium bromide complexed to ct-DNA is shown in Fig. 8. The spectra are in accord with literature data,[31] but, slight differences were observed compared with the control experiments with the large-volume Couette *LD* cell (see below).

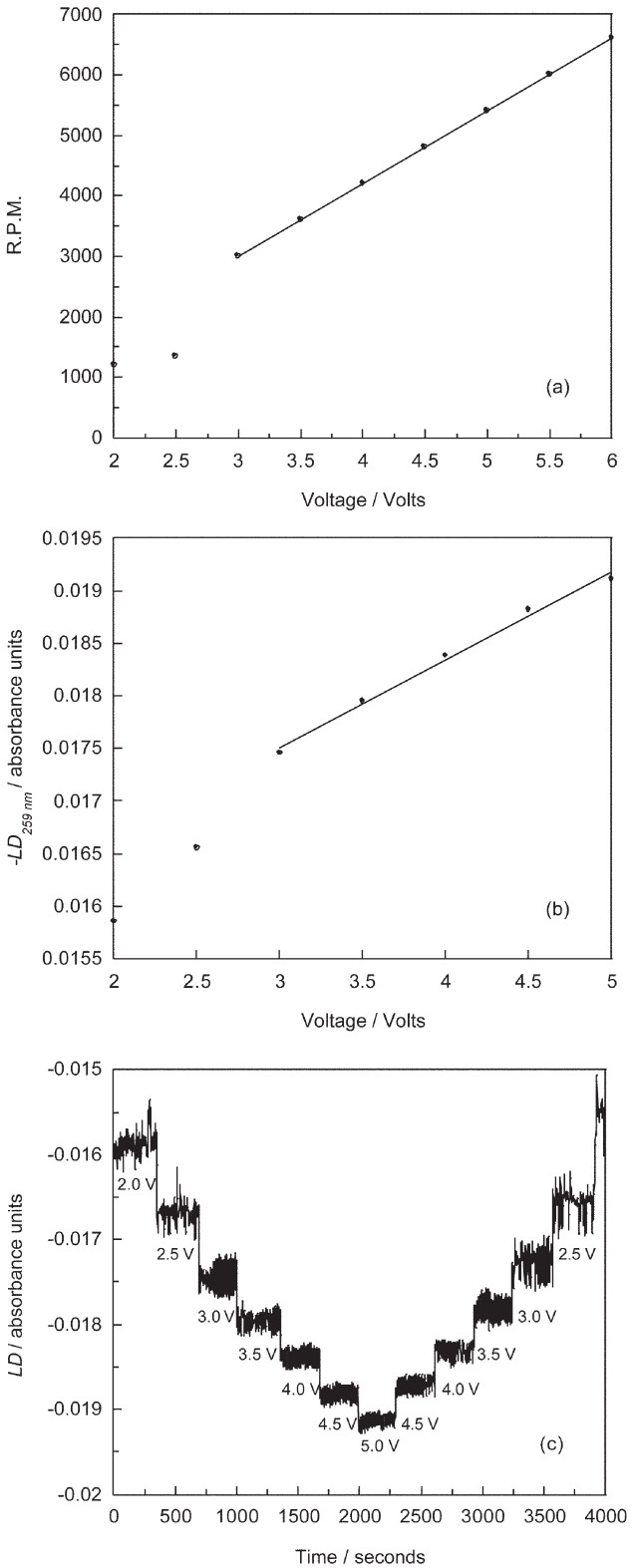


Fig. 6 (a) The relationship between voltage and rpm of the microvolume Couette *LD* cell, (b) the relationship between LD_{259} and voltage and (c) *LD* intensity variation at different voltages. All experiments were carried out using ct-DNA (~200 μM).

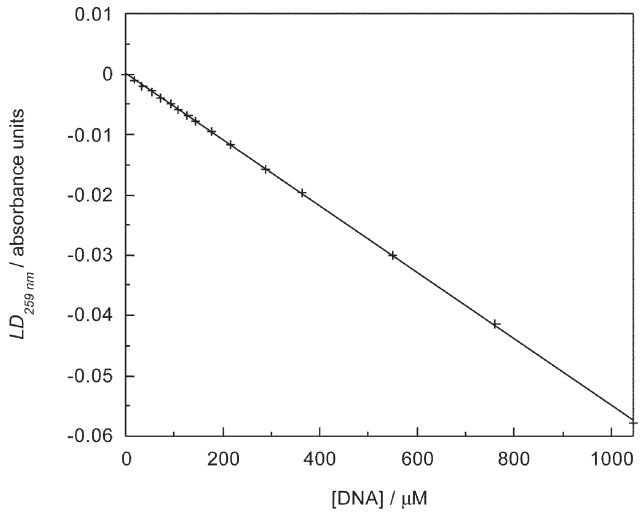


Fig. 7 LD_{259} *versus* DNA concentration in the microvolume Couette *LD* cell at 4 V.

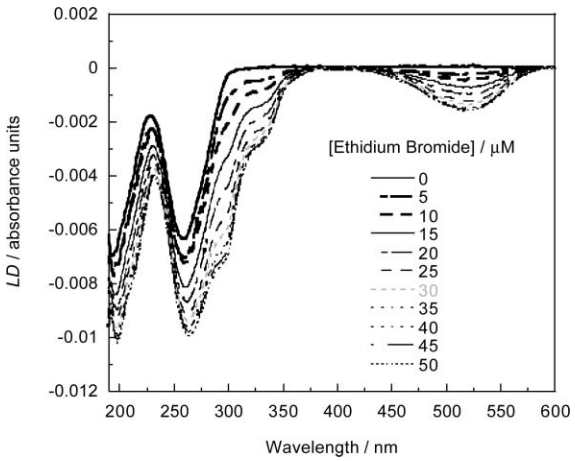


Fig. 8 *LD* spectra of ct-DNA (200 μM) and different concentrations of ethidium bromide (0–50 μM) in sodium cacodylate buffer (10 mM) pH 7.0 and NaCl (10 mM).

Capillary cell *versus* large-volume cell for DNA

The microvolume Couette cell was evaluated against LV Cells 1 and 2 to investigate cell to cell reproducibility. LD_{259} as a function of DNA concentration was determined for the capillary cell and LV Cell 2 (Fig. 9). The microvolume Couette *LD* cell follows the Beer–Lambert law somewhat better than LV Cell 2 (but its signal is smaller as it has a smaller path length and also smaller orientation effect than LV Cell 2 (though not LV Cell 1)). LV Cell 1, which is the older of the two LV cells, deviates from the capillary cell at low wavelength.

DNA and ethidium bromide cell to cell comparison

Fig. 10 shows that it is not possible to simply rescale the microvolume Couette cell data to directly overlay with the LV Cell 2 for DNA–ethidium bromide (*i.e.* over a wide wavelength range). This was found also to be the case for different ratios of DNA : ethidium bromide. In order to determine whether this was an effect of the different shear forces being applied to the samples, the shear force on both cells was calculated following the model for shear force within small annular gaps given by Nordén *et al.*[10]

$$G \approx \frac{\omega r_o}{r_o - r_i} \quad (2)$$

where G is the shear force, ω the angular velocity and r_o and r_i the radii of the outer and inner cylinders respectively. The velocity profile is shown in Fig. 11.

The microvolume Couette cell, assuming 3 V (minimum voltage required for linearity of DNA signal) is equivalent to 3000 rpm, *i.e.* 310 rad s^{-1}, and r_o and r_i are 1.45 and 1.2 mm respectively, resulting in a shear force of $\sim$1800 s^{-1}. An approximately equivalent shear force in LV Cell 2 requires ω = 70.1 rad s^{-1} using values of r_o and r_i of 13 and 12.5 mm; this equates to an rpm value of 670 rpm and a voltage of 1.62 V. LD^r data were therefore collected for both systems

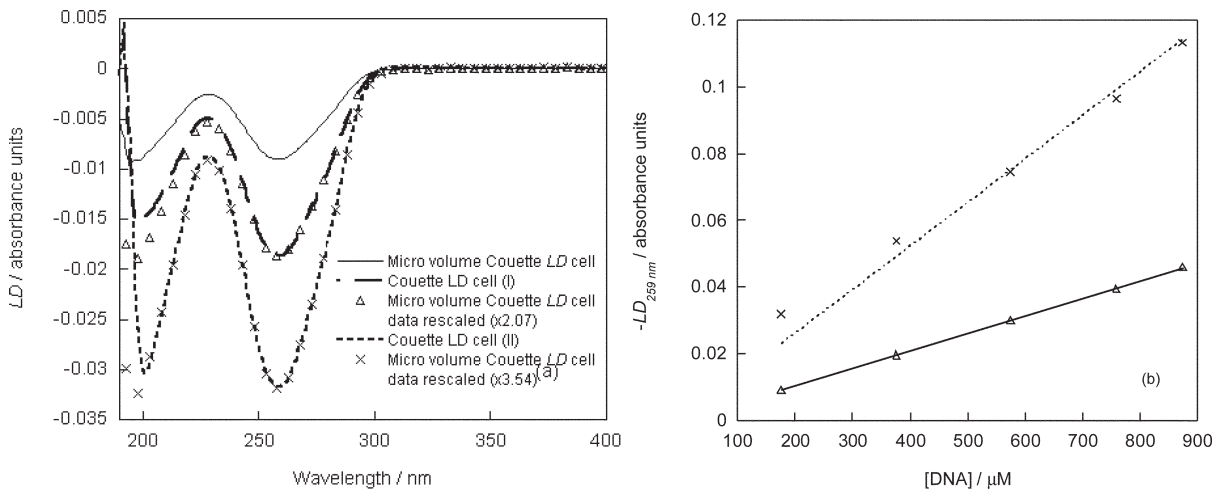


Fig. 9 *LD* of DNA (176 μM). (a) Wavelength scans in microvolume Couette *LD* cell (—), LV Cell 1 (– – – –) and LV Cell 2 (- - -), with rescaling of the microvolume Couette *LD* cell spectrum to overlay at 259 nm on LV Cell 1 (△ △ △) and LV Cell 2 (× × ×) spectra; (b) 259 nm magnitudes for microvolume Couette LD cell (solid line) and LV Cell 2 (dotted line).

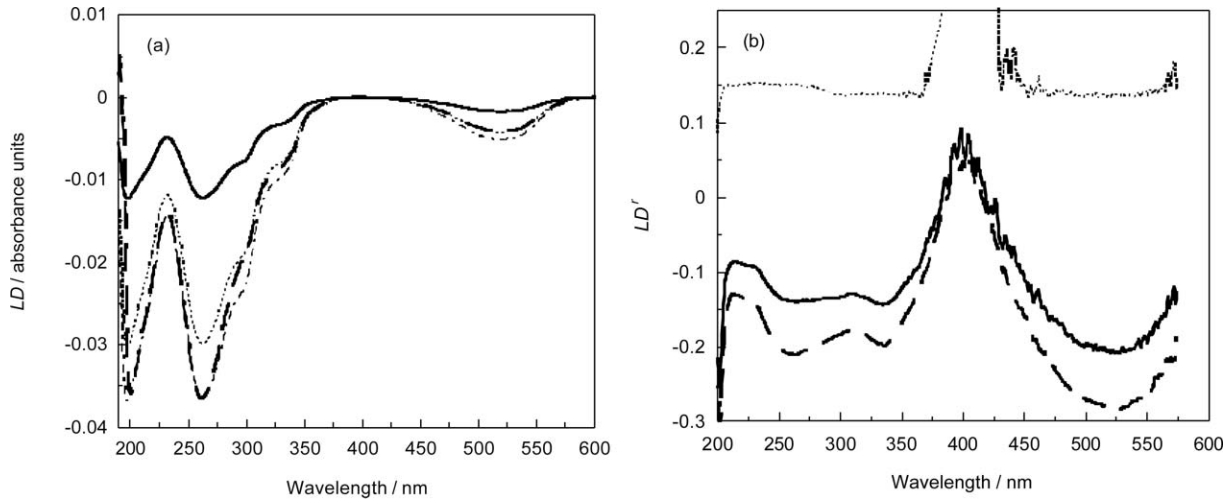


Fig. 10 (a) *LD* spectra of DNA (176 μM) and ethidium bromide (50 μM) in NaCl (10 mM) and sodium cacodylate buffer (10 mM) pH 7.0 in the microvolume Couette cell (—) and LV Cell 2 (– – –). The *LD* spectrum for the microvolume Couette *LD* cell has been rescaled to match the signal of LV Cell 2 at 259 nm (- — -) and at 520 nm (·····). (b) LD^r spectra of *LD* spectra in part (a) and the ratio of Cell 2 to capillary cell *LD* (divided by 10) (·····).

using similar shear forces (Fig. 10). The spectra still show the same differences, so the discrepancy is not a function of shear force. The LD^r spectra and the ratio between the LD^r's of the two cells are constant in the DNA region and constant across the ethidium region—but with, relatively, the microvolume cell having the smaller signal for DNA or the larger for ethidium. At lower ethidium loadings the same effect is observed, but in addition the long-wavelength ethidium band has a slightly lower ratio than the shorter wavelength bands. Another key difference between the two cells is the fact that LV has an inner rotating cylinder while the microvolume cell has an outer rotating one. The former has a more complicated flow behaviour as summarised in Fig. 12 and discussed below.

The most definitive work on Couette systems after the time of Mallock[17] and Couette[15] was that of Taylor who investigated the stability of a viscous flow with two cylinders rotating in the same direction and in opposite directions and later went on to conclude that if the outer cylinder only rotates, the flow is more stable than if only the inner cylinder rotates.[16,19,21] Taylor distinguished two broad categories of fluid states by their symmetry under rotation and reflection: laminar flow and turbulent flow. Laminar flow (also known as Couette flow) is the most stable flow and usually occurs at low flow velocities. It is this state that is required for *LD*. In turbulent flow, vortices, eddies and wakes make the flow unpredictable. Between laminar and turbulent flow a state called Taylor-vortex flow exists made up of helical vortices. A useful dimensionless variable to summarise the relative significance of the viscous effect compared to the inertia effect is given by the Reynolds number (R_e) (named after Osbourne Reynolds[32,33]). R_e, in the form applicable for a fluid in the annulus between two concentric cylinders, is defined

$$R_e = \frac{\omega_1 r_1 \delta}{\nu} \tag{3}$$

where ω_1 is the angular velocity of one rotating cylinder in rad s^{-1}, r_1 is the radius of that rotating cylinder in m, δ is the gap width $|r_2 - r_1|$ in m, and ν the kinematic viscosity in $m^2\ s^{-1}$ (kinematic viscosity is equal to absolute or dynamic viscosity divided by density).[34] A complex transition diagram has been published by Andereck *et al.*[36] in 1986 (Fig. 12) to summarise when each state is expected.

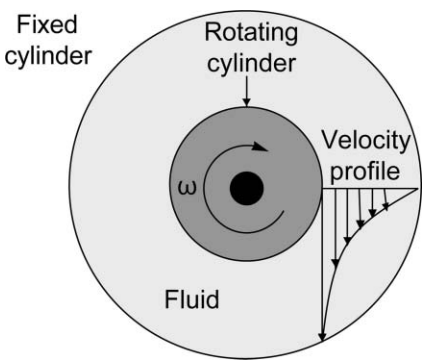


Fig. 11 Plan of a Couette cell showing the velocity profile of a fluid contained between concentric cylinders as the inner cylinder rotates, redrawn from Wilkes.[35] The arrows represent the direction of the flow at the point where the axes touch the inner cylinder and the lengths represent the velocity.

The significance of this diagram for the microvolume and LV Cells is summarised in Table 2 if we assume: (1) the solution is or resembles water, so has a kinematic viscosity[34] of $10^{-6}\ m^2\ s^{-1}$ and (2) DNA may be represented as an infinite length cylinder. In particular, at least in principle, the outer rotating cell should remain in the Couette flow domain whereas the inner rotating cell is close to the regime of Taylor-vortex flow. In such a flow state the fluid forms cylindrical vortices about the flow direction. This would mean that the DNA is oriented in toroids about the flow direction, with the regions of ethidium intercalation orienting more effectively. Somewhat perversely, the DNA molecules so oriented will reduce the orientation parameter for ethidium relative to the DNA bases (in Couette flow it is the other way

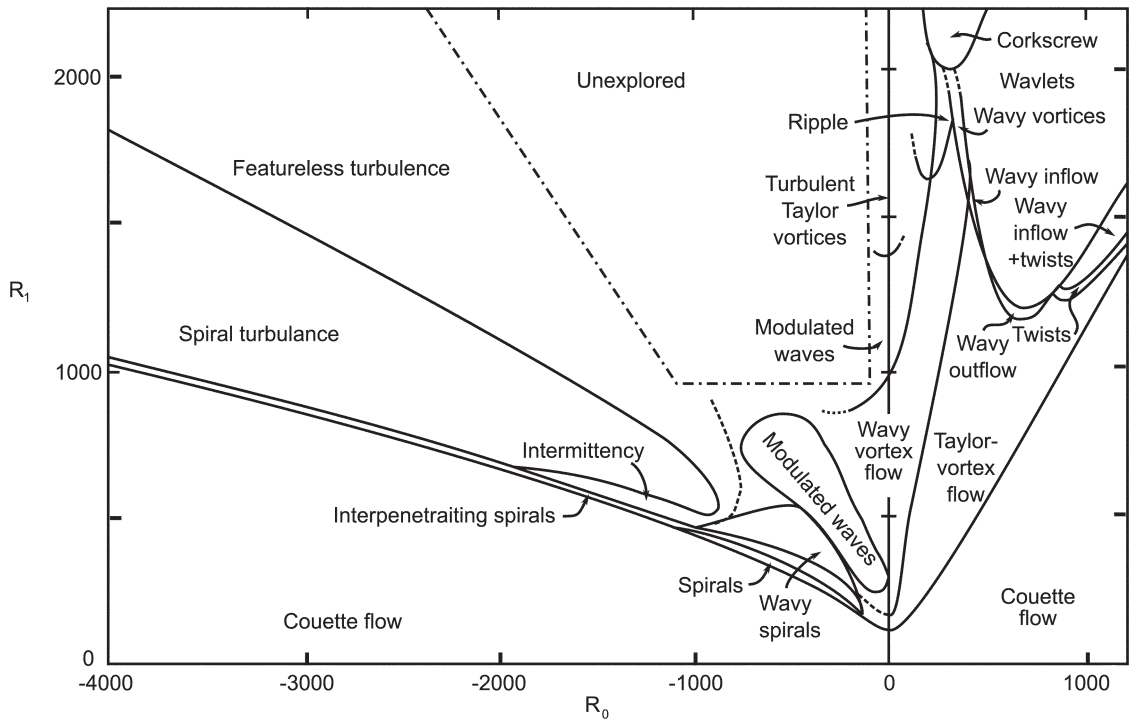


Fig. 12 Diagram of Taylor–Couette flow states from that of Andereck *et al.*,[36] where R_i and R_o are the inner and outer cylinder Reynolds numbers respectively.

Table 2 Comparison of fluid flow in the LV Cell 2 and the microvolume Couette *LD* cell for solutions of the same kinematic viscosity as that of water

		ω				
	High/low speed of rotation	/rpm	/rad s^{-1}	R_o	R_i	Flow state
Couette *LD* cell	Low (1 V)	300	31.4	0	196.25	Couette flow–Taylor-vortex flow
$r_1 = 0.0125$ m	High (4 V)	2000	209.4	0	1308.75	Turbulent flow
$\delta = 0.0005$ m						
Microvolume Couette *LD* cell	Low (3 V)	3000	314.2	113.90	0	Couette flow
$r_1 = 0.00145$ m	High (6 V)	6600	691.2	250.5	0	Couette flow
$\delta = 0.00025$ m						

round since the DNA is stiffer), resulting in relative reduction of the orientation parameter for ethidium-bound DNA *versus* free DNA.

Conclusions

The aim of the work reported in this paper has been to establish validation protocols for *LD* and in particular to determine to what extent the data collected in the newly designed microvolume capillary *LD* cells can be trusted. The most important result is that the new configuration with lenses to focus the light beam on the capillary does not cause any artefacts in the observed spectrum compared with older cells where this is not required. The data from the microvolume cell follow the Beer–Lambert law and have signals linear in voltage after a threshold value. The data are reproducible, with only $\sim$1% variation with careful reloading of samples, and can be used over a wide bandwidth range. The data from previously designed large-volume cells are consistent with those from the capillary for pure DNA samples. However, for DNA with ethidium bromide intercalated in some DNA binding sites the spectra show different orientation parameters in the DNA base and bound ethidium regions. This proved not to be due to differences in flow velocities, but rather to be intrinsic features of the designs of the cells. The microvolume cells have an outer rotating capillary and inner stationary rod, so their fluid flow is more stable and the data more reliable from the microvolume cells than those from inner rotating large-volume cells that have been previously used in our laboratory. At least in principle, the capillary cells maintain Couette flow which means we are indeed orienting the samples along the direction of flow, in contrast to the LV cells where it seems that some element of Taylor-vortex flow is operating even at low voltages. However, it should be noted that in practice bubbles readily form in the capillary cells at high voltages in samples containing detergents. Additional advantages of the microvolume cells are that it has been possible to design a thermostatted jacket for the cell for steady state temperature work, it requires only $\sim$25 μL sample (with the size of Araldite plug we choose to use), though $\sim$40 μL makes sample loading

easier. It should be noted, however, that the large-volume cells have an undeniable advantage when titration experiments need to be undertaken.

Acknowledgements

The collaboration of Crystal Precision Optics, Rugby UK has been a key factor in the design and building of the micro-volume capillary cells and is gratefully acknowledged. Syngenta UK (PhD studentship to RM), the BBSRC (REI20503), The Association of Clinical Biochemists, and the EPSRC (GR/T09224/01) are also thanked for financial support for this work. Jim MacDonald is acknowledged for drawing Fig. 12. The referees' careful work has also been very valuable.

References

1 T. R. Dafforn, D. J. Halsall and A. Rodger, *Chem. Commun.*, 2001, 2410–2411.
2 T. R. Dafforn, J. Rajendra, D. J. Halsall, L. C. Serpell and A. Rodger, *Biophys. J.*, 2004, **86**, 404–410.
3 T. R. Dafforn and A. Rodger, *Curr. Opin. Struct. Biol.*, 2004, **14**, 541–546.
4 T. Yanagida, M. Taniguchi and F. Oosawa, *J. Mol. Biol.*, 1974, **90**, 509–522.
5 M. Taniguchi and R. Kuriyama, *Biochim. Biophys. Acta*, 1978, **533**, 538–541.
6 A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez and M. J. Hannon, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4051–4057.
7 A. Rodger and B. Nordén, *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, UK, 1997.
8 M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chem.*, 2001, **40**, 879–884.
9 B. Nordén, *Appl. Spectrosc. Rev.*, 1978, **14**, 157–248.
10 B. Nordén, M. Kubista and T. Kurucsev, *Q. Rev. Biophys.*, 1992, **25**, 51–170.
11 P. Chou and W. C. Johnson, *J. Am. Chem. Soc.*, 1993, **115**, 1205–1214.
12 A. Rodger, *Methods Enzymol.*, 1993, **226**, 232–258.
13 L. B. A. Johansson and A. Davidsson, Analysis and Application of Linear Dichroism on Membranes—Description of a Linear-Dichroism Spectrometer, *J. Chem. Soc., Faraday Trans. 1*, 1985, **81**, 1375–1388.
14 R. Marrington, T. R. Dafforn, D. J. Halsall and A. Rodger, *Biophys. J.*, 2004, **87**, 2002–2012.
15 M. Couette, Etudes sur le frottement des liquides, *Ann. Chim. Phys.*, 1890, **6**, 433–510.
16 R. J. Donnelly, *Phys. Today*, 1991, 32–39.
17 A. Mallock, *Proc. R. Soc. London*, 1888, **45**, 126–132.
18 A. Mallock, *Philos. Trans. R. Soc. London, Ser. A*, 1896, **187**, 41–56.
19 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1923, **223**, 289–343.
20 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1936, **157**, 546–564.
21 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1936, **157**, 565–578.
22 A. Wada and S. Kozawa, *J. Polym. Sci., Part A*, 1964, **2**, 853–864.
23 S. Higashi, M. Kasai, F. Oosawa and A. Wada, *J. Mol. Biol.*, 1963, **7**, 421–430.
24 J. Hofricheter and W. Eaton, *Annu.Rev. Biophys. Bioeng.*, 1976, **5**, 511–560.
25 F. Oosawa, Y. Maeda, S. Fujime, S. Ishiwata, T. Yanagida and M. Taniguchi, *J. Mechanochem. Cell Motil.*, 1977, **4**, 63–78.
26 A. Wada, *Appl. Spectrosc. Rev.*, 1972, **6**, 1–30.
27 A. Wada, *Biopolymers*, 1964, **2**, 361–380.
28 M. Taniguchi, A. Yamaguchi and T. Taniguchi, *Biochim. Biophys. Acta*, 1971, **251**, 164–171.
29 R. Marrington, E. Small, A. Rodger, T. R. Dafforn and S. Addinall, *J. Biol. Chem.*, 2004, **47**, 48821–4882.
30 R. D. Wells, J. E. Larson, R. C. Grant, B. E. Shortle and C. R. Cantor, *J. Mol. Biol.*, 1970, **54**, 465–497.
31 B. Nordén and F. Tjerneld, *Biophys. Chem.*, 1976, **4**, 191–198.
32 O. Reynolds, *Proc. R. Soc. London*, 1883, **35**, 84–99.
33 O. Reynolds, *Philos. Trans. R. Soc. London*, 1883, **174**, 935–982.
34 M. C. Potter and D. C. Wiggert, *Mechanics of Fluids*, Brooks/Cole Thomson Learning, Pacific Grove, CA, 2002.
35 J. O. Wilkes, *Fluid Mechanics for Chemical Engineers*, Prentice Hall PTR, Upper Saddle River, NJ, 1999.
36 C. D. Andereck, S. S. Liu and H. L. Swinney, *J. Fluid Mech.*, 1986, **164**, 155–183.

Job: Nienhaus 372-6
Chapter: 17/Rodger

Compositor: NF
Date: 12/04
Revision: 2nd Pass

17

Circular Dichroism Spectroscopy for the Study of Protein–Ligand Interactions

Alison Rodger, Rachel Marrington, David Roper, and Stuart Windsor

Summary

Circular dichroism (CD) is the difference in absorption of left and right circularly polarized light, usually by a solution containing the molecules of interest. A signal is only measured for chiral molecules such as proteins. A CD spectrum provides information about the bonds and structures responsible for this chirality. When a small molecule (or ligand) binds to a protein, it acquires an induced CD (ICD) spectrum through chiral perturbation to its structure or electron rearrangements. The wavelengths of this ICD are determined by the ligand's own absorption spectrum, and the intensity of the ICD spectrum is determined by the strength and geometry of its interaction with the protein. Thus, ICD can be used to probe the binding of ligands to proteins. This chapter outlines protein CD and ICD, together with some of the issues relating to experimental design and implementation.

Key Words: Circular dichroism; proteins; chirality; ligand binding; induced circular dichroism.

1. Introduction

A key feature of any biological system is its chirality or asymmetry or handedness: a chiral molecule has a mirror image that is not superposable on itself. This means the two mirror images cannot be rotated so that they look exactly the same. Macroscopic as well as smaller scale chirality is ultimately dependent on the molecular level. Because many molecules in biological systems are chiral and are present in only one enantiomeric form, the macroscopic structures they build are also chiral. Molecular chirality is perhaps most obvious with a helical molecule such as the double helical structure of B–DNA, but it is true of all proteins and nucleic acids.

From: *Methods in Molecular Biology, vol. 305: Protein–Ligand Interactions: Methods and Applications*
Edited by: G. U. Nienhaus © Humana Press Inc., Totowa, NJ

Circular dichroism (CD), which is the difference in absorption of left and right circularly polarized light, is probably the simplest technique for non-destructively providing solution phase structural information about chiral molecules. Many ligands (usually small molecules that bind to a macromolecule) in biological systems are also chiral, in which case they have their own CD spectrum, which will probably be perturbed when the ligands bind to a protein. If a ligand is achiral, then it will have no intrinsic CD but will gain an induced CD (ICD) signal in its transitions when it binds to a protein. It is this ICD signal that contains the information about the asymmetry of the protein–ligand interaction. In this chapter, we will focus on how to measure CD spectra of proteins and protein–ligand complexes and how to analyze the data.

1.1. Protein Absorbance Spectroscopy

In order to understand CD spectroscopy and use the data intelligently, it is essential that one measures the absorbance spectrum of one's sample since this shows where to expect CD signals. The Beer–Lambert law for the absorption of light by a sample of concentration C is

$$A = \varepsilon C \ell \tag{1}$$

where ℓ is the length of the sample through which the light passes, and ε is known as the molar extinction coefficient and depends on the wavelength at which the absorbance is being measured. If ℓ is measured in cm and C in M = mol dm^{-3}, then ε has units of mol^{-1} dm^3 cm^{-1}. The Beer–Lambert law is valid as long as the spectrometer can measure the intensity of photons passing through the sample (i.e., the concentration is not so large that essentially all photons are absorbed), *and* there are no concentration-dependent intermolecular interactions.

In the case of peptides and proteins, the spectroscopy of the amide bonds, the side chains, and any prosthetic groups (such as haems) determines the observed UV/visible absorption spectrum with their intensities and wavelengths often being affected by the local environment of the groups. UV spectra of proteins are usually divided into the *near* and *far* UV regions. The near UV in this context means 250–300 nm and is often described as the aromatic region because of the absorption of the aromatic amino acids, though transitions of disulfide bonds (cysteine–cysteine bonds) also contribute to the total absorption intensity in this region. The far UV (< 250 nm) is dominated by absorption as a result of the peptide backbone of the protein (*see* **Fig. 1**), but transitions from some side chains also contribute. The far UV absorbance of a protein is typically of the order of a factor of 100 stronger than the near UV absorbance.

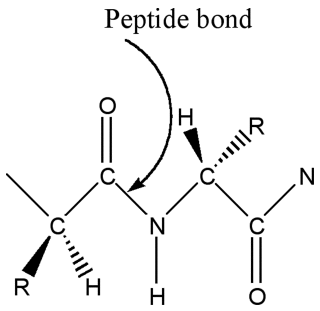


Fig. 1. Peptide units of amino acids linked to form a peptide backbone. R denotes side chains.

1.2. Protein Circular Dichroism

CD may be defined as the difference in absorption of left and right circularly polarized light:

$$CD = \Delta A = A_l - A_r \tag{2}$$

where the subscripts indicate the polarization of light incident on the sample. The CD of macromolecules differs from that of small molecules in that the spectroscopy of the component units (amino acids, in the case of proteins) is at least partly understood: the form of the CD spectra result largely from the arrangement of the units in space, and the coupling between the transitions of those units. The CD of proteins is most commonly used in one of two ways: (i) to probe the secondary structure of the protein itself, and (ii) to probe its binding interaction with small molecules (usually referred to as ligands). Even though both of these applications are in practice largely empirical, it is useful to have some understanding of the origin of the signal being measured.

Proteins are long chains of amino acids. If the side chain on an amino acid is not H (i.e., not glycine), then the tetrahedral carbon of that amino acid is a chiral center and we should expect a CD signal in transitions of the neighboring amide groups and side chains. If, however, there is free rotation about the bonds of the main chain, the observed amide $n \rightarrow \pi^*$ (a weak transition with $\varepsilon \sim 100$ mol^{-1} dm^3 cm^{-1} occurring at 210–230 nm) and $\pi \rightarrow \pi^*$ (a stronger transition with $\varepsilon \sim 7000$ mol^{-1} dm^3 cm^{-1} centered at 190–200 nm) ***(1,2)*** CD will be relatively small. Thus, a truly random or denatured protein will have only a small CD signal at the wavelengths accessible to most CD machines. The most stable conformation of a protein under physiological conditions, however, is not random but composed of well-defined structures that give macro-chiral units (including the well established α-helix and β-sheet) with significant CD intensities that add together to give the observed spectrum. A wide range of

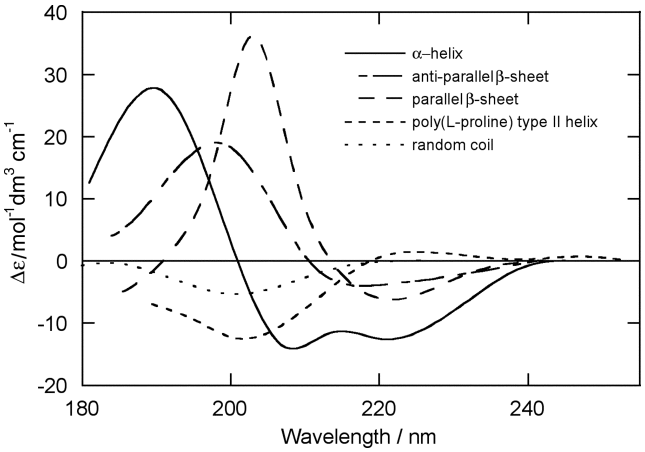


Fig. 2. Backbone (i.e., 180–250 nm) CD spectra for different protein secondary structural motifs ***(14,15)***. α-helix (solid line, average of poly(γ-methyl-L-glutamate) in F_6iPrOH and poly(L-alanine) as a film), antiparallel β-sheet (dashed line, BOC (L-alanine)7OMe as a film), parallel β-sheet (dash dot dot line, BOC (L-valine)7OMe as a film), poly(L-proline) II type helix (dotted line, systemin in 10 m*M* phosphate buffer at 5°C, pH 7.0), and random coil (dash dot line, collagen at 45°C in 0.01 *M* phosphate buffer, pH 3.5).

fitting programs are available to deconvolute the experimental data into percentages of the structural motifs ***(3)***. Some work in terms of the spectra of a set of identified structural motifs, others use a basis set of real protein spectra where the percentage of different structural units is known. The approach of Johnson et al. ***(4)*** uses a large basis set of such proteins, chooses a subset of these randomly until a good fit is obtained. Almost all approaches will give a good estimate of α-helix content *if* the protein concentration is known (*see* **Note 1**). This is because the α-helix profile is the most distinct and also largest in magnitude as illustrated in **Fig. 2**. The distinctive shape arises from the coupling of the $\pi \rightarrow \pi^*$ transition moments in each amide chromophore and results in a component at about 208 nm which contributes to the characteristic α-helix CD spectrum. Good estimates will also be obtained for proteins without α-helices. Reasonable estimates will be obtained for mixed structure proteins, though here accurate values of the amino acid residue concentration are essential.

1.3. Units of Circular Dichroism Spectroscopy

If equal intensities of left and right handed circularly polarized light (i.e., linearly polarized light) were passed through a CD active sample at the same time, the result of differential absorption would be slightly elliptically polar-

ized light. For historical reasons many CD spectropolarimeters therefore produce a CD spectrum in units describing this degree of *ellipticity*, θ, in millidegrees even though they measure differential absorbance (Δ*A*). The conversion between these two is ***(3)***:

$$CD = \Delta A = \frac{4\pi\theta\,(\mathrm{deg})}{180\ln 10} = \frac{\theta\,(\mathrm{mdeg})}{32,982} \tag{3}$$

The CD Analog of the Beer–Lambert Law,

$$CD = \Delta A = \Delta\varepsilon C\ell \tag{4}$$

AU: ok to change $\Delta\varepsilon cl$ to $\Delta\varepsilon C\ell$ as already defined in Eq. 1?

where Δε is the differential molar extinction coefficient (in units of $mol^{-1}\ dm^3\ cm^{-1}$), *C* is concentration in mol dm^{-3}, and ℓ is the path length of sample through which the light passes, in cm. This allows conversion to molar absorbance (or if millidegrees have been retained, molar ellipticity). In this context, it is important to have clear in one's mind what concentration is being used. For example, protein concentrations are sometimes in mg/mL, sometimes in moles of protein molecules per dm^3, and sometimes (as for protein structure fitting programmes) in moles of amino acid residues per dm^3.

2. Materials

2.1. Samples and Solvents

The materials for protein–ligand CD spectroscopy are most obviously the protein and ligand samples of interest. In addition, there is the solvent in which the sample is prepared. For proteins this usually means a buffer, as many proteins need buffers to retain their structure. However, if the buffer of choice has an absorbance above 1 at the concentrations to be used anywhere in the wavelength range of interest for CD, then the buffer will be absorbing most of the light and the quality of your CD spectrum will be poor or even rubbish. Phosphate is a good spectroscopy buffer; Tris is satisfactory; cacodylate is spectroscopically good but is an arsenic salt; acetate is far from ideal and may limit your accessible range to above approx 215 nm; and any buffers with aromatic groups are impossible. It must be remembered that the solvent and buffer can sometimes affect protein structure, and therefore the CD spectra obtained. Even if there are no structural changes, changing the chemical environment can affect electronic transitions, and therefore also affect the CD spectra.

2.2. Instrument Calibration

It is important to calibrate the CD machine for both intensity and wavelength. It is straightforward to carry out a check on the state of calibration of a CD instrument. If the intensity calibration is outside specification, it is usually

easy to reset the intensity yourself, but it is more difficult to reset the wavelength, and a service engineer should be called.

Wavelength calibration (*see* **Note 2**) requires a wavelength standard. One option is a solid neodymium filter, which usually comes with the CD spectrometer. The absorbance maximum of this should be 586 ± 0.8 nm. Another is a Holmium Oxide filter, or solution, which has a characteristic spectrum across the near UV and visible regions. The absorbance maximum in a CD machine can usually be determined by determining where the gain voltage on the photomultiplier tube (often called the high tension voltage) is maximum.

Intensity calibration (*see* **Note 3**) is conventionally carried out using aqueous ammonium d-10-camphor sulfonate (ACS), which has been related to the established hydroscopic primary standard camphor sulfonic acid (CSA) ***(5,6)***. ACS is available from a number of suppliers (including Katayama Chemical Co. 05-1251), although care should be taken to ensure its enantiomeric as well as the chemical purity. The concentration of ACS usually used is 0.06% w/v, i.e., 60.00 mg in 100 mL or 6.000 mg in 10 mL, which should give a CD signal of 190.4 ± 1 mdeg at 290.5 nm (the CD maximum). (Note: CSA has a different molecular weight from ACS so a 0.06% solution has different molar concentrations, and hence different CD signals for the two standards.) It is important to weigh the solid material accurately (to 4 significant figures). For 6 mg this requires a 6 figure balance. The standard solution can be stored at 4°C. However, the storage time seems to be dependent on the container in which it is stored and storage of longer than 2 wk may be problematic.

2.3. Path Length Calibration

In order to use and compare CD data, it is important that the path length of the cell is measured—often the nominal path length of a cell is far from the actual path length. This is particularly important for short path length cells often used in the far UV (secondary structure) region. One approach is to use an aqueous potassium chromate solution of known concentration (*see* **Note 4**). The Beer–Lambert Law can then be used to calculate the path length from a measurement of absorbance. For 0.01 mm path length, a 0.2 *M* potassium chromate solution is required. For longer path lengths, a more dilute solution is required. To prepare a 0.2 *M* potassium chromate solution, accurately weigh 0.971 g potassium chromate (Aldrich 21,661-5, 99% ACS reagent) and transfer it to a 25 mL volumetric flask. Add approx 20 mL water and one pellet of potassium hydroxide (e.g., AnalaR BDH 102104V). Make the solution up to volume and mix well. This solution may be stored indefinitely at 4°C. If lower concentration samples are prepared by diluting the 0.2 *M* solution the potassium hydroxide is also diluted, and so the solution stability is reduced.

3. Methods

3.1. Collecting Circular Dichroism Data

Measuring a CD spectrum is a routine procedure assuming one has access to a CD spectropolarimeter. If your sample gives a good UV-visible absorbance spectrum then it is highly likely that (if it is chiral) you will get a good CD spectrum. The essential features of a CD spectrometer are a source of (more or less) monochromatic left and right circularly polarized light and a means of detecting the difference in absorbance of the two polarizations of light. CD machines are much more expensive than UV machines, as one is typically expecting absorbance differences of the order of 10^{-3} to 10^{-4} to be accurately measured against a background absorbance of just less than 1. For protein CD the spectropolarimeters also need to be nitrogen purged, not only to prevent the high energy lamps producing ozone and therefore damaging the delicate optics within the instrument, but also to avoid having O_2 in the sample compartment absorbing the incident radiation and limiting the lowest wavelengths that can be measured. In practice this means a moderate nitrogen flow rate (3–5 cm^3/min) at all times, with an increase to 20 or more cm^3/min when collecting data below 190 nm. **Subheading 3.1.1.** describes the important steps you need to consider while making CD measurements

3.1.1. Before You Start

In order to collect good quality and reliable CD data you should ensure that your instrument is properly calibrated both in terms of wavelength (*see* **Note 2**) and intensity (*see* **Note 3**). You should also carefully select a sample cell that is suitable for CD measurements (*see* **Note 5**), and is an appropriate path length for the concentration of the sample that you are measuring (*see* **Note 6**). Because the actual path length of cells can differ significantly from the nominal values, you should always ensure that you calibrate the path length of the cell that you are using (*see* **Note 4**). For fixed path length sealed cells, it is useful to calibrate all of your cells, and keep a calibration table handy for easy reference.

3.1.2. Sample Absorbance

Because the wavelength and absorbance ranges are the same for both normal absorption and CD it is advisable to run a normal absorption spectrum of the sample for which you wish to measure the CD spectrum first. *Always* leave the reference beam of the absorbance spectrometer empty for this experiment, because we are concerned with the total absorbance of the sample including any achiral molecules in the buffer. It is preferable to have the maximum absorbance less than 1.5 (and certainly less than 2).

3.1.3. Wavelength Range

CD spectrometers usually scan from longer wavelengths to shorter ones. Ensure you collect data for at least 20 nm to the longer wavelength side of any absorbance band. Select a wavelength range starting so that there is at least 20 nm of zero absorbance beyond the normal absorption envelope(s) of interest. When the baseline spectrum (*see* **Subheading 3.1.4.**) is subtracted from the sample spectrum, the region outside the absorption envelope should be flat. If it is not then this probably means either there is a very weak absorbance band that has a large dissymmetry factor, and hence large CD signal compared with its normal absorbance intensity, or, more probably, there is some light scattering by the sample. Sources of light scattering include dirty cuvets (inside or outside), undissolved sample, condensation of samples, and particulate samples (*see* **Note 7**). Given that the wavelengths of light being scattered are less than 1 μm, one does not necessarily expect to be able to see by eye the presence of such particles.

3.1.4. Baseline and Zeroing

A machine baseline (i.e., CD spectrum of air) measured on a standard CD spectrometer will not be flat as the optical components in the instrument are birefringent. This can be ignored by storing a machine baseline in the instrument. However, the cuvet used for an experiment will also have its own CD spectrum—CD matched cuvets usually have slightly different intrinsic spectra. So always collect a baseline spectrum of your solvent/buffer under the same conditions as the sample spectrum using the same cuvet in the same orientation with respect to the light beam. Subtract the baseline spectrum from the sample spectrum to produce the final CD plot. Although one can often have the cuvet baseline automatically subtracted by the software, this may not be a good idea for small signals or sticky samples that could be retained on the cuvet—it is better to see any problems in the baseline spectrum. It is often the case that even when the baseline is subtracted, the CD is not exactly zero outside the absorption envelope. However, if it is flat outside the absorption envelope, the spectrum may be zeroed by adding or subtracting a constant (either within the spectrometer software or using your chosen data plotting software).

3.1.5. Signal-to-Noise Ratio

The signal-to-noise ratio in a CD spectrum increases with: $\sqrt{n}$, where n is the number of times the spectrum is accumulated (and the data averaged); $\sqrt{t}$, where t is the time over which the machine averages each data point; and $\sqrt{I}$ where I is the intensity of the light beam. The actual definition of t depends on whether your instrument moves to a wavelength and sits there to collect data or whether it continuously scans. The best way to deal with this issue is to try some scans

with different time constants or response times to see if the spectrum is distorted. I is usually influenced by bandwidth and also wavelength, because the lamp does not have uniform intensity at all wavelengths. In fact, when spectra become noisy one of the reasons is often that the lamp is old and its light intensity decreased. Most CD spectrometers have both short timescale (millisecond to minutes) and long timescale (minutes to hours) baseline variations. To average over short timescale fluctuations collect a number of spectra or have a longer response time. Longer timescale fluctuations are usually dealt with by alternating collection of sample and baseline spectra or by assuming the fluctuation is a wavelength independent shift of the zero point. While appearing to be wishful thinking, the assumption of linear baseline drift seems to be valid especially with more recent instruments.

3.1.6. Parameter Sets

CD spectrometers give the operator considerable control over response time (τ), scan speed (s), bandwidth (b) [the wavelength range (error) of the incident light], and data interval (d). To optimize signal-to-noise effects (*see* **Subheading 3.1.5.**) t should be selected to be as large as possible subject approximately to

$$\tau \times s \leq \frac{b}{2}$$

If t is too long for the chosen s and b, then maxima of peaks (both positive and negative) will be cut off and their wavelengths shifted. A control scan using

$$\tau' = \frac{\tau}{2} \text{ (or } s' = 2s)$$

should be used to check that spectra are not being distorted by the chosen parameters. The data interval determines the wavelength interval between data points. This may or may not have implications for the time a scan takes depending on whether the instrument operates in a stepped scan mode or continuous scan mode.

Scan speeds of 50 nm/min, t = 1 s, b = 1 nm, and a data step of 0.5 nm seem to be a good starting point as a parameter set for most protein–ligand experiments where the samples have the broad band shapes usually found for protein samples and most ligands. It is often advisable to perform a fast preliminary scan to determine whether there is any point in collecting an accurate spectrum. The most significant consideration for reliable data is whether the sample has too high an absorbance—this is summarized by the high-tension voltage (or equivalent label) of the photomultiplier tube. It may also be necessary to select an appropriate sensitivity scale—the need for this will be indicated by a flat straight line across peaks in your spectrum.

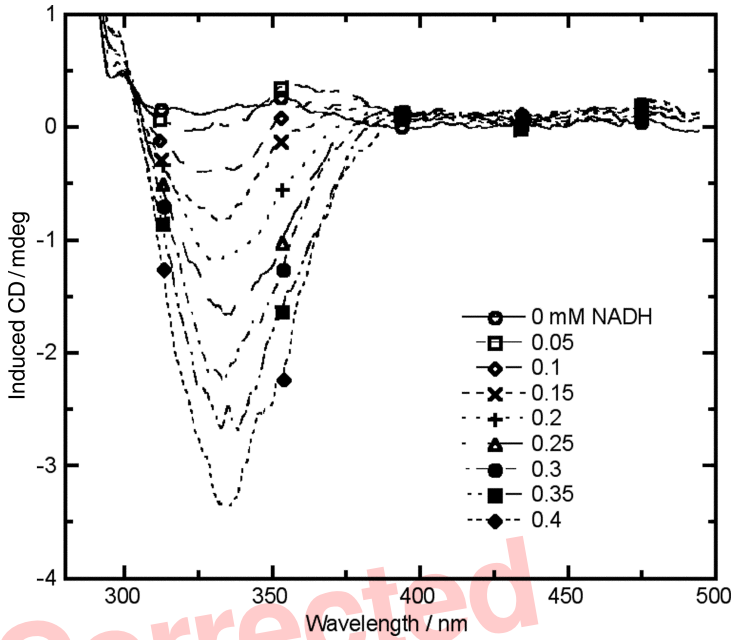


Fig. 3. CD of lactate dehydrogenase (LDH, 22.5 μ*M*) with increasing concentrations of NADH. LDH concentration was kept constant by adding equal volumes of NADH (from a 10 m*M* stock solution) and LDH from a 45 μ*M* stock solution. The measured spectrum is the sum of the intrinsic spectra of the two components, so there is no evidence of interaction by CD.

3.1.7. Linearity

It is good practice to check that the CD signal is proportional to sample concentration by running a spectrum of a sample and a diluted sample and checking that the signal scales with concentration. If it does not then you have evidence of solute–solute interactions in your sample or too high an absorbance for the instrument to be able to cope with.

3.2. Protein–Ligand Interactions Probed by Circular Dichroism: Titration Experiments

When a ligand binds to a protein molecule, its transitions will acquire an ICD spectrum. If it is itself achiral then this will change its CD spectrum from zero to something; if it is chiral then its ICD is the difference between its own CD spectrum and the spectrum in the presence of the protein. **Figure 3** shows a titration series where it appears that a large ICD is being observed when NADH binds to lactate dehydrogenase. However, the ligand is chiral and we

are only measuring the ligand's intrinsic CD. The protein will also acquire an ICD (though with a large protein it may only be a very small change). The ICD in the protein regions of the spectrum are usually a combination of protein and ligand ICD as most ligands have absorption intensity below 300 nm. A number of examples are given in **Subheading 3.3.**

If there is just one binding mode (or equal occupancy of more than one) between protein and ligand then the magnitude of the ICD will be proportional to the concentration of protein–ligand complex (*see* **Subheading 3.3.**). The challenge is to determine the proportionality constant, especially if high or low loadings cannot be achieved experimentally. If there is a succession of binding modes being occupied, the ICD spectral shape usually changes.

Apart from the challenges of analyzing the data once collected, there are a number of issues that must be considered for ligand binding experiments. One typically proceeds by titrating (or adding) one solution to another. There are an infinite number of ways of doing this. Conceptually, the simplest is to make up a series of independent solutions. However, this is usually the least accurate method as pipeting errors are most significant. Alternatively, one can add very small volumes of a highly concentrated stock solution of ligand or protein, and either ignore the dilution effect or account for it assuming it has no structural effect on the protein. A simple way to avoid dilution effects is as follows. Consider a starting sample that has concentration x *M* of species *X*. Each time y µL of *Y* is added, also add y µL of a $2x$ *M* solution of *X*. The concentration of *X* remains constant at x M. An infinite number of variations on this theme are possible, one of which is to add a solution already containing x *M* of *X*. One must always consider whether the order of addition of components or the stock solution concentrations (particularly local high concentrations during addition of the titrant) in such an experiment has any effect on the nature of the interaction.

Whether the protein or the ligand concentration is kept constant during a titration series depends on a number of issues. If the protein CD spectrum changes as a function of its own concentration then it is usually wise to keep it constant. However, if it does not, by keeping the ligand concentration constant one can more easily monitor any changes in ligand ICD as a function of protein:ligand ratio. Any ICD in the protein region of the spectrum is likely to be composed of both protein and ligand ICD spectra as it is hard to find ligands without any spectroscopy below 300 nm.

3.3. Examples of Protein-Ligand Circular Dichroism Spectra

3.3.1. Protein–Solvent Interactions

An illustration is given in **Fig. 4** of the effect on a protein CD spectrum of a ligand with no spectroscopy. In this case solvent induced conformational

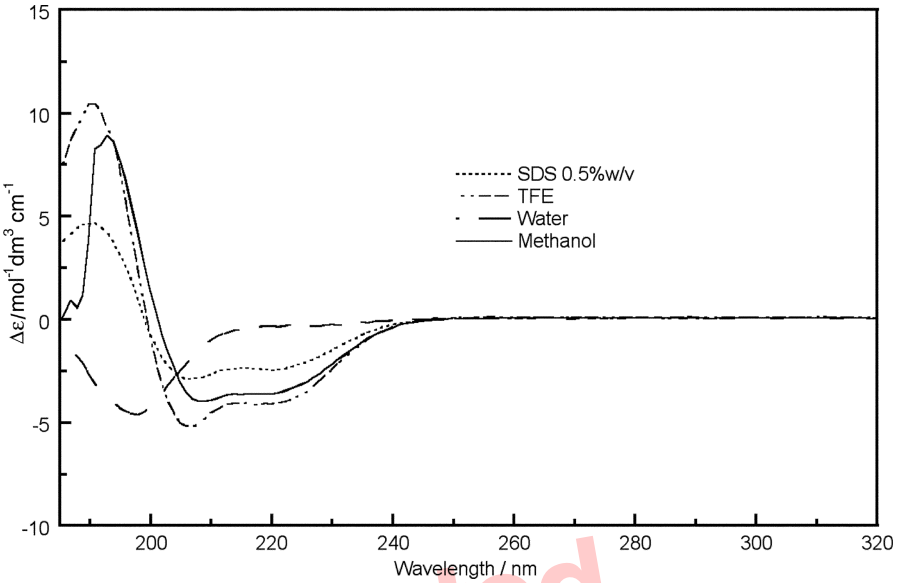


Fig. 4. CD of a peptide (MSLSRRQFIQASGIALCAGAVPLKASA) in different solvents ***(16)***. Water shows a random coil structure. The TFE spectrum is approx 50% α-helical.

changes in a short peptide (MSLSRRQFIQASGIALCAGAVPLKASA in single letter code) are illustrated. A change from approx 50% α-helical (in TFE) to random coil (in water) is observed—this is an extreme change and not to be expected for most ligand binding.

3.3.2. Protein–Protein Interactions

In many instances, the ligand that binds to a protein is in fact another protein. **Figure 5** shows the CD spectra measured for F-actin (the polymerized form of actin), a myosin molecule (S1) and the mixture of the two. The theoretical spectrum calculated from adding the actin and myosin spectra is also shown. The difference between this theoretical spectrum and the experimental mixture is small but clear and reproducible, showing that either the actin or the myosin secondary structure is perturbed by their interaction or there is some electronic perturbation of their transitions. It is not possible to distinguish these possibilities.

3.3.3. ICD of a Protein–Ligand System

Bovine serum albumin binds almost anything that can be put into solution with it, however, one of its main roles is as a steroid transporter protein in

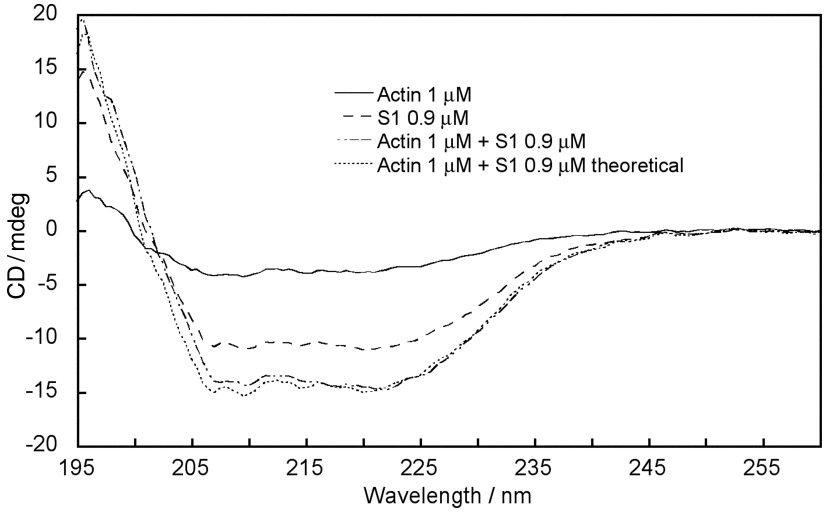


Fig. 5. CD spectra of F-actin (1 μ*M* in actin monomer), S1 myosin (0.9 μ*M*), the mixture of the two proteins and the sum of the two independent spectra. All samples were made up in 5 m*M* Tris buffer, pH 8.0, 20 μ*M* $CaCl_2$, 1 m*M* ATP, 1 m*M* $MgCl_2$, and 50 m*M* KCl. Pathelength used was 1 mm.

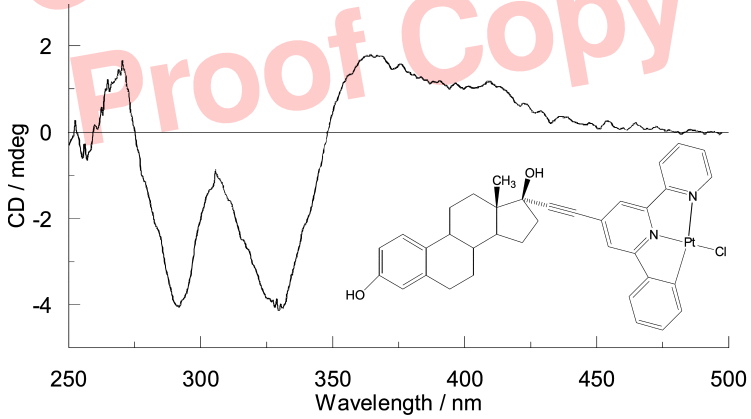


Fig. 6. ICD for PtEEtpy (30 μ*M*, structure is shown in the figure) bound to bovine serum albumin (30 μ*M* in 5 m*M* sodium cacodylate buffer).

blood. An example of the CD it induces in a steroid coupled to a platinum metal complex is shown in **Fig. 6**. The signal above 300 nm is caused by the perturbations to the ligand spectroscopy. That below 300 nm is a combination of protein and ligand ICD signals.

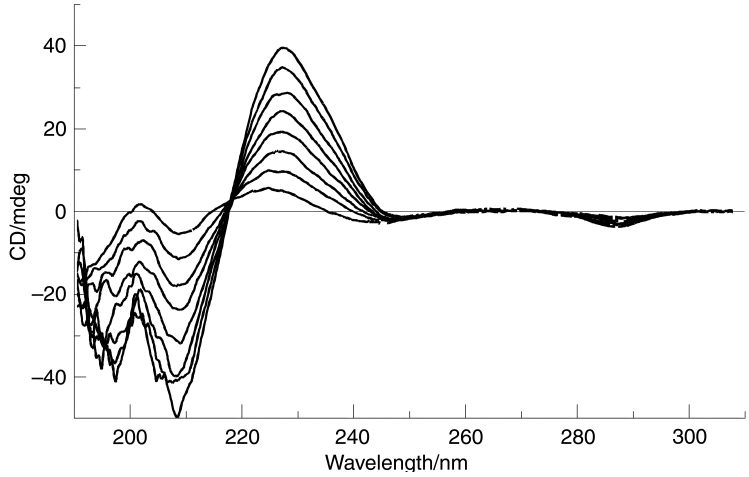


Fig. 7. ICD measured when increasing volumes of a concentrated solution of vancomycin containing 25 μ*M* ristocetin are added to a 25 μ*M* ristocetin solution. The intrinsic CD spectra of the peptides have been subtracted to give only the ICD. The 0 μ*M* vancomycin spectrum is therefore a flat line and has been omitted. Vancomycin concentrations (increasing ICD magnitude) are: 25 μ*M*, 50 μ*M*, 75 μ*M*, 100 μ*M*, 125 μ*M*, 150 μ*M*, 175 μ*M*, 200 μ*M*. (Data from **ref. *17***).

3.3.4. Peptide–Peptide Titration Series

Vancomycin and ristocetin are glycopeptide antibiotics that prevent cross-linking and transglycosylation during bacterial cell wall formation. Noncovalent dimerization plays a key role in their activity, and CD can be used to give binding constants for such an interaction. The data in **Fig. 7** are for the equilibrium

$$V + R \rightleftharpoons V - R$$

where V denotes vancomycin, R ristocetin and V–R their noncovalent complex. The ICD is assumed to be proportional to the concentration of dimer. **Figure 7** shows the data collected as increasing concentrations of vancomycin are added to a solution where the concentration of the ristocetin is maintained at 25 μ*M*. The equilibrium constant for the dimerization can then be calculated to be $K_{dimerisation} = 20 \pm 5\ (mM)^{-1}$ (*see* **Note 8** for methods for determining equilibrium binding constants).

3.4. Other Techniques

CD measurements of protein–ligand binding are in fact not always sensitive enough to detect ligand binding, if there is little change in the protein second-

ary structure. The near UV CD is only perturbed if the ligand binds near an aromatic residue; and the ligand ICD signal may be very small. Thus, ligand binding may be more effectively probed by other techniques such as fluorescence or even absorbance spectroscopy. Measuring fluorescence polarization anisotropy (FPA) may in fact achieve some of the advantages of CD spectroscopy in that it is the difference between two polarizations of radiation so much of the background signal gets subtracted off.

CD is particularly suited to probing chiral molecules because achiral effects are canceled out. Vibrational circular dichroism (VCD), Raman optical acitivity (ROA), and optical rotatory dispersion (ORD) do the same thing. VCD and ROA are attractive as they probe the vibrations of the molecule, which can usually be readily assigned and there may well be many more transitions than in UV–CD providing complementary information. Both VCD and ROA currently require much higher sample concentrations and generally much longer data accumulations than UV–CD, so they are not nearly as widely available. ORD is related to CD by the Kramers–Kronig transformation, so in principle contains the same information. However, the information is much harder to extract and data over the full wavelength range is at least in principle required.

For fibrous proteins such as actin (*see* **Subheading 3.3.2.**) the binding of a ligand is usually readily probed by flow oriented linear dichroism ***(7)*** (the difference in absorption of light linearly polarized along the orientation axis and perpendicular to that axis) because the ligand is invisible until it is oriented by the fiber to which it binds.

4. Notes

1. Any quantitative application of CD spectroscopy to protein structure fitting or ligand binding needs a fairly accurate estimate of protein concentration. For a well-known protein the extinction coefficient at 280 nm may well be available on one of the many protein data base websites or in the literature. It is then a fairly simple matter to take a sample (probably of approx 1 mg/mL), measure its absorbance spectrum in a 1 cm cuvet, and use the Beer–Lambert law. However, you should note that the extinction coefficient from a database is almost certainly a theoretical one determined from adding the contributions of aromatic amino acids and disulfide bonds. Thus, it ignores environmental effects. At the very least one should completely denature one's protein to use the theoretical value, then determine a native ε value for later use. Denaturing the protein is usually carried out with high concentrations of guanidinium chloride, whose purity may be sufficiently suspect that it should be determined using refractive index measurements ***(8)***. Alternatively, for a totally new protein, amino acid analysis may be recommended ***(9)***.
2. After the CD machine has been on for more than 30 min run a spectrum of a neodymium filter from 610 nm to 560 nm with instrument parameters set for a

fairly slow scan with small data pitch, for example: 0.1 nm data pitch, 20 nm/min scan speed, 1 accumulation, 0.25 s response time, 1.0 nm band width. Note, the photomultiplier tube voltage—the absorbance maximum is also the HT voltage maximum and it should occur at 586 ± 0.8 nm. It is a good idea to note the wavelength accuracy in an equipment maintenance log. If the wavelength accuracy is not within specification, but the shift is constant across the wavelength range (check for the same variation with ACS, *see* below), then you can recalibrate the spectrum accordingly. However, it is advisable to call in an engineer.

3. The intensity calibration of a CD machine is usually carried out by collecting a spectrum from 350 nm to 250 nm of 0.06% aqueous ammonium d-10-camphor sulfonate (ACS) in a 1 cm path length cuvet. A typical set of instrument parameters is: 0.1 nm data pitch, 50 nm/min scan speed, 1 accumulation, 1 s response time, 1.0 nm band width. Subtract a water baseline run with the same cuvet and parameters. The wavelength and intensity of the peak should be 190.4 ± 1 mdeg at 290.5 nm. It is a good idea to record the values in the instrument log. If the intensity is not within stated limits, use an independently made fresh ACS standard and repeat the calibration test. If the value is reproducible all subsequent data may be scaled to bring the intensity to the correct value. Alternatively, the instrument may be recalibrated.

   This approach to instrument calibration is based on the assumption that a single point calibration is sufficient for the whole spectrum. This is transparently not the case and work is in progress to try to establish a single solution calibrant for the full wavelength range. Of particular concern is the instrument performance below 200 nm where light scattering effects become significant as the lamp intensity is reduced. The 0.06% ACS solution in a 1 mm cuvet can be used to monitor this as the ACS has a negative CD maximum at 191 nm. The magnitude of the ratio of the negative to positive peaks is probably approx 2.1. Certainly, a value below 2.0 indicates poor instrument performance ***(10)***.

4. It is important to determine the cuvet path length if a short path length (less than 1 mm) is being used. One millimeter and longer path length cuvets can usually be assumed to be that specified by the manufacturer (though if pathlength is important this should be confirmed). 0.01 mm path lengths are almost never close to that specified, and indeed the path length of a filled demountable cuvet varies from fill to fill and user to user.

   Short path lengths (0.1 mm and shorter) can be determined using interference fringes on an empty cell. However, for the shorter path length demountable cuvets how a given operator fills the cuvet will affect the path length. The best method we have found is for each user to fill the demountable cell with the potassium chromate solution of appropriate concentration and measure the UV/visible absorption spectrum from 600 nm to 350 nm. The path length is then calculated using the Beer–Lambert law ($\varepsilon = 4830$ mol^{-1} dm^3 cm^{-1} at 372 nm). It is important to always assemble the cell the same way (mark the cuvet at one end with a pencil and note which edge is the beveled edge). Path lengths of demountable cuvets do vary over time (as the edges of the cell get worn). It is important

for at least three measurements to be performed. A new user of 0.01 mm cuvets, in our experience, takes hours of reloading and remeasuring to obtain a reproducible path length.

5. Either cylindrical or rectangular cuvets may be used for CD. Cylindrical cells are usually deemed to have lower birefringence (baseline CD) than rectangular cuvets, however, if UV and CD *matching* is requested when the cuvets are purchased rectangular cuvets seem to be equally good. Water-jacketed cylindrical cells enable the sample to be thermostatted most simply, and also take the least sample volume for a given path length. With these cuvets, you must check that the configuration of your light beam and cuvet holder to ensure that the light beam passes through the sample and not the quartz walls and cooling water parts of the cuvet. Rectangular cells have a number of advantages over cylindrical cuvets for the 1 mm and longer path length experiments: they are cheaper, may be used in standard absorption spectrophotometers (so CD and normal absorption data may be collected on exactly the same sample), and may be used for a protein-ligand serial titration experiment as approx 60% of a rectangular cell can be empty for the first spectrum and gradually filled.

   If path lengths of 0.1 mm or less are required it is probably best to use demountable cuvets where the sample is dropped onto a quartz disk or plate that is etched to a predefined depth, and then another quartz disk/plate is carefully placed on top. Titrations are not possible in demountable cuvets unless independent samples are made.

   All of the light beam incident upon the cuvet must pass through the sample and not be clipped or reflected by the walls or base of the cell or the meniscus of the solution, otherwise the measured spectrum is affected by scattered light. Thus, the narrow cells often used to minimize sample volume in a normal absorption spectrophotometers *cannot* be used for CD unless the light beam is chopped or focused or is intrinsically small. While focusing of the light beam is possible, one must ensure that (i) the lenses used for the focusing are not themselves significantly birefringent (CD active), (ii) the light beam does not diverge and hit the sides of the cuvet while passing through the sample, and (iii) the whole light beam incident on the sample is collected by the photomultiplier tube (PMT). The light beam must not be focused too tightly on the PMT itself, otherwise the PMT may be damaged.

   For UV/visible (*see* above note) CD, high-quality quartz cuvets that transmit the full wavelength range of UV/visible (*see* above note) light are required. In the visible region glass may be used, but it is generally advisable to use quartz even here. Plastic cuvets typically have high intrinsic birefringence so should be avoided. In any case, the need to run a baseline of each cuvet used (*see* below) removes the usual attraction of disposable plastic cuvets.
6. The required path length for the protein backbone region spectrum (from 260 nm to 190 nm) may be estimated on the basis that a 1 mm cuvet probably requires a approx 0.1–0.2 mg/mL protein solution. Sometimes it is desirable to adjust concentrations to use an available cuvet, sometimes it is desirable to choose a path

length to avoid dilution of a sample (e.g., monoclonal antibodies usually have a slightly concentration dependent CD spectrum). The path length required for the aromatic region (from 300 nm to 250 nm) depends on the concentration of aromatic chromophores in the protein. For a protein with no aromatic groups and no disulfide bonds, there will be no aromatic region CD signal whatever the concentration or path length used. Typically, 100 times more protein needs to be in the light beam for near UV measurements than for far UV measurements. The path length required to measure the CD induced into ligand transitions upon binding to a protein is chosen to give an absorbance of approx 1 at the wavelength of interest (usually around the absorption peaks of the ligand).

7. It is essential that the cuvet is cleaned well. Any deposit of chiral material on the quartz will have a CD spectrum. Sometimes one just hopes this subtracts off with a baseline, however, this is not good practice. To clean a cuvet one may proceed as follows. Rinse it well at least three times with high purity water (18.2 MΩ) followed by ethanol or acetone. Dry the inside of a nondemountable cuvet with nitrogen or compressed air (but beware of oil deposits from the compressor) or a hairdryer. Dry the outside of a cuvet with a tissue, wipe with a lens cloth, and remove any fibers with a nitrogen line. If the cuvet shows traces of protein residue (as most easily shown by a protein CD spectrum being observed for the baseline), wash well with detergent (e.g., Hellmanex), and rinse with water. If the residue still remains, place the cuvet in a solution of 6 *M* nitric acid (beware of local safety issues here, e.g., acetone and nitric acid are explosive) or Hellmanex (make sure the cleaning agent gets inside the cuvet) and allow it to stand for 10 min or longer before removing and rinsing thoroughly with water.
8. When measuring protein–ligand interactions, one is frequently interested in the strength of the interaction. The simplest measure of the binding strength between a protein and a ligand is the equilibrium binding constant,

$$K = \frac{L_b}{L_f S_f} \tag{5}$$

for the equilibrium

$$\text{free ligand} + \text{empty protein binding site} \rightleftharpoons \text{protein} - \text{ligand complex} \tag{6}$$

where L_b is the concentration of bound ligand, L_f is the concentration of free ligands, and S_f is the free site concentration. The total binding site concentration is

$$S_{tot} = nC_M \tag{7}$$

where CM is the macromolecule concentration and n is the number of sites per protein. For proteins, it is usually safe to assume $n = 1$ for the first binding site. CD can be used to determine binding constants if the strength is of the order of $(\mu M)^{-1}$ or a little weaker or stronger simply because of the concentration of the samples required to get a reasonable signal. If no ICD is observed then it is likely that the interaction is weaker. If the ICD is proportional to concentration of ligand (or protein) added then it is too high to be measured by CD.

Assuming that CD is a useful method for determining K, methods are outlined below which are particularly appropriate for use with CD data where the starting point for equilibrium constant determination is usually:

$$L_b = \alpha\rho \tag{8}$$

where ρ is the CD signal at a chosen wavelength and a (which is a function of wavelength) is a constant over the range of binding ratios being considered. The simplest means of determining a is usually from the low binding ratio limit where all the ligand is assumed to be bound so L_f is assumed to be zero. If this is indeed the case, then a plot of ρ vs L (total ligand concentration) at constant protein concentration should be a straight line with slope a. Alternatively, the maximum ICD signal may be used to determine α if n is known (in this limit $L_b = C_M$). One can then calculate K from a point-by-point analysis of the data. Alternatively, the Scatchard plot is probably still the most widely-used method of combining data from a titration series.

The method most widely used in one form or another for determining K is the one developed by Scatchard ***(11)***. **Equation 5** is rearranged as follows:

$$\frac{r}{L_f} = \frac{KS_f}{C_M} = Kn - rK \tag{9}$$

where

$$r = \frac{L_b}{C_M} \tag{10}$$

because $S_f = S_{tot} - L_b$. So, a plot of r/L_f vs r has slope $-K$ and y-intercept K/n. The x-intercept occurs where $r = n$. L_b, and hence L_f, may be determined directly from the CD if a has been determined as discussed earlier.

9. It is not always possible to get data for either high- or low-binding ratio limits to determine a. In such cases the intrinsic method is used:

   Equation 5 may be written ***(12)***

$$K = \frac{\alpha\rho}{(S_{tot} - \alpha\rho)(L_{tot} - \alpha\rho)} \tag{11}$$

which can be rearranged to give

$$L_{tot} = \frac{L_{tot}S_{tot}}{\alpha\rho} - S_{tot} + \alpha\rho - \frac{1}{K} \tag{12}$$

for two different total ligand concentrations, and the same protein site concentration, i.e.,

$$\frac{L_{tot}^k - L_{tot}^j}{\rho^k - \rho^j} = \frac{S_{tot}}{\alpha}\left(\frac{\frac{L_{tot}^k}{\rho^k} - \frac{L_{tot}^j}{\rho^j}}{\rho^k - \rho^j}\right) + \alpha \tag{13}$$

Thus, a plot of

$$y = \frac{L_{tot}^k - L_{tot}^j}{\rho^k - \rho^j} \quad \text{vs} \quad x = \left(\frac{\frac{L_{tot}^k}{\rho^k} - \frac{L_{tot}^j}{\rho^j}}{\rho^k - \rho^j} \right)$$

should be a straight line with slope $C_M(\alpha/n)^{-1}$ and intercept α. The concentration of bound molecules in any sample may then be determined as may n. The equilibrium binding constant, K, may then be calculated. Alternatively, using these accurate values of n and a Scatchard plot (*see* previously) may be used to determine the best value of K using all the data points.

It is sometimes convenient to perform experiments with constant ligand and varying macromolecule concentration. In this case C_M, and hence S_{tot}, are the variables and L_{tot} is fixed. Rather than **Eq. 13** we then use:

$$\frac{S_{tot}^k - S_{tot}^j}{\rho^k - \rho^j} = \frac{L_{tot}}{\alpha} \left(\frac{\frac{S_{tot}^k}{\rho^k} - \frac{S_{tot}^j}{\rho^j}}{\rho^k - \rho^j} \right) + \alpha \tag{14}$$

or equivalently

$$\frac{C_M^k - S_M^j}{\rho^k - \rho^j} = \frac{L_{tot}}{\alpha} \left(\frac{\frac{C_M^k}{\rho^k} - \frac{C_M^j}{\rho^j}}{\rho^k - \rho^j} \right) + \alpha / n \tag{15}$$

10. A glance at almost any reference dealing with titrations will make one realize that the simple equilibrium model we have assumed is quite probably invalid and also the options for data analysis are almost endless ***(13)***. In particular, whereas the traditional approaches mentioned previously involve linearizing the data in some way as this enables one to see by eye whether such an approach is valid, with available computers and packages it is fairly simple to convert almost any model into a plot and determine constants for it without weighting data in any way.

References

1. Nakanishi, K., Berova, N., and Woody, R. W. (ed.) (1994) *Circular Dichroism: Principles and Applications.* VCH, New York.
2. Rodger, A. and Nordén, B (1997) *Circular and Linear Dichroism.* Oxford University Press, Oxford.
3. Johnson, W. C. Fitting programs are available to deconvolute the experimental data into percentages of the structural motifs. Website: e.g., http://www.cryst.bbk.ac.uk/cdweb/html/home.html; Dicroweb: a facility of the BBSRC Centre for Protein and Membrane Structure and Dynamics. http://oregonstate.edu/dept/biochem/faculty/johnson.html. Date accessed: November 20, 2004.
4. Johnson, W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins*: *Structure, Function and Genetics* **7,** 307–312.

5. Chen, G. C. and Yang, J. T. (1977) Two-point calibration of circular dichrometer with D-10-camphorsulfonic acid. *Anal. Lett.* **10,** 1195–1207.
6. Takakuwa, T., Konno, T., and Meguro, H. (1985) A new standard substance for calibration of circular dichroism: ammonium D-10-camphorsulfonate. *Anal. Sci.* **1,** 215–218.
7. Dafforn, T. R., Halsall, D. J., Serpell, L. C., Rajendra, J., and Rodger, A. (2004) The use of linear dichroism to determine the orientation of secondary structural elements within protein fibres. *Biophys. J.* **86,** 404–410.
8. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves *Methods Enzymol.* **131,** 266–280.
9. Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.
10. Miles, A. J., Wien, F., Lees, J. G., Rodger, A., Janes, R. W., and Wallace, B. A. (2003) Calibration and standardisation of synchrotron radiation circular dichroism and conventional circular dichroism spectrophotometers. *Spectroscopy* **17,** 653–661.
11. Scatchard, G. (1949) The attraction of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.* **51,** 660–672.
12. Rodger, A. (1993) Linear dichroism. *Methods Enzymol.* **226,** 232–258.
13. Polster, J. and Lachman, H. (1989) *Spectrometric Titrations: Analysis of Chemical Equilibria.* VCH Verlagsgesellschaft, Weinheim, Germany.
14. Johnson, W. C., Jr. (1988) Secondary structure of proteins through circular dichroism spectroscopy. *Ann. Rev. Biophys. Biophys. Chem.* **17,** 145–166.
15. Johnson, W. C., Jr. (1985) Circular dichroism and its empirical application to biopolymers. *Methods Biochem. Anal.* **31,** 61–163.
16. Miguel, M. S., Marrington, R., Rodger, P. M., Rodger, A., and Robinson, C. (2003) An *Escherichia coli* twin-arginine signal peptide switches between helical and unstructured conformations depending on hydrophobicity of the environment. *Euro. J. Biochem.* **270,** 1–8.
17. Green, P. (1999) PhD Thesis, University of Warwick, UK.

Photochemistry and Photobiology, 2006, 82: 20–30

Spectroscopic Identification of Binding Modes of Anthracene Probes and DNA Sequence Recognition†

Willy B. Tan[1], Akhilesh Bhambhani[1], Michael R. Duff[1], Alison Rodger[2] and Challa V. Kumar*[1]

[1]Department of Chemistry, University of Connecticut, Storrs, CT
[2]Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK

Received 24 May 2005; accepted 22 September 2005; published online 29 September 2005 DOI: 10.1562/2005-05-24-RA-539

ABSTRACT

The binding properties of two anthracene derivatives with calf thymus DNA (CT DNA), poly(dA-dT), and poly(dG)·poly(dC) are reported. One contained bulky, cyclic cationic substituents at the 9 and 10 positions, and the other carried acylic, branched, cationic substituents. Binding of the probes to the DNA was examined by calorimetry, spectroscopy and helix melting studies. The cyclic derivative indicated exothermic binding, strong hypochromism, bathochromism, positive induced circular dichroism (CD, 300–400 nm), significant unwinding of the helix, large increases in the helix melting temperature, strong but negative linear dichroism (LD, 300–400 nm) and considerable stabilization of the helix. In contrast, the acyclic analog indicated thermoneutral binding, smaller hypochromism, no bathochromism, very weak induced CD, and no change in the helix melting temperature with any of the DNA polymers. A sharp distinction between the binding properties of the two probes is indicated, and both have intrinsic binding constants of ~10^6 M^{-1} for the three polymers. However, when the ionic strength of the medium was lowered (10 m*M* NaCl), the absorption as well as CD spectral changes associated with the binding of the acyclic derivative corresponded with those of the cyclic derivative. The acyclic derivative showed large preference (10-fold) for poly(dG)·poly(dC) over poly(dA-dT), whereas the cyclic analog showed no preference. The characteristic spectroscopic signatures of the two distinct binding modes of these probes will be helpful in deciphering the interaction of other anthracene derivatives with DNA.

INTRODUCTION

The interaction of small ligands with DNA is important in many biological processes (1). Several anthracene derivatives, such as pseudourea, for example, show antitumor activity (2), and their interaction with DNA was thought to be necessary for biological activity (3). Better understanding of these interactions will be useful for the rational design of ligands to target specific regions of DNA and for enhanced therapeutic activity. One approach to achieve this goal is to systematically vary the structural features of the ligand, and investigate how these variations influence its DNA binding properties (4,5).

The chemical reactivity of the anthracene nucleus allows for the decoration of this ligand with a variety of substituents at the 9 and 10 positions. When the anthracene ring is intercalated with its long axis parallel to the base pair long axis, these substituents are expected to project into the grooves for interaction with the helix. This approach provides opportunities in controlling the nature, type, and magnitude of the interactions with the DNA helix. The low polarizability, good absorption in the near-UV region, strong fluorescence, and a long-lived triplet excited state of the anthracene system are additional attractive features. Photoactivation of the anthracene probes, for example, has been shown to cause DNA strand cleavage with high selectivity (6). The planar, anthracene ring system is suitable for intercalation into the DNA helix, and many anthryl probes have been suggested to prefer intercalation (7–11). The linear ring system is expected to support good overlap with the DNA base pairs. These features provoked a systematic investigation of the DNA binding properties of selected anthracene derivatives.

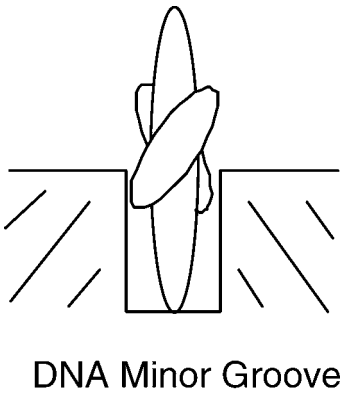
One of the key issues in understanding the interaction of ligands with DNA is the identification of the binding mode or modes by which the ligand binds to DNA. Two of the dominant DNA binding modes of anthracene derivatives are intercalation and groove binding, which are shown in Chart 1. The steric, electrostatic and geometric requirements of these modes are long recognized to be different (12), but the exact role of substituents in controlling the relative populations of these modes is not clear. Small changes in the relative energies of these modes will result in large differences in their respective populations. This is because the population of the upper state decreases exponentially with increase in the energy gap between the states, as dictated by the Boltzmann distribution (13).

The relative energies of the binding modes may be fine-tuned by placing appropriate substituents on the anthracene ring system. Two anthracene derivatives carrying branched substituents (BPA and BDEA, Chart 2) are chosen to test this hypothesis. Substantially large substituents on the intercalator are expected to inhibit intercalation due to steric blockage (14,15). Data are presented here that demonstrate that even small differences in the substituents can cause large changes in the binding preferences.

* To whom correspondence should be addressed: Department of Chemistry, University of Connecticut, Storrs, CT 06269-3060, USA. Fax: 860-486-2981; e-mail: challa.kumar@uconn.edu

† This paper is part of a special issue dedicated to Professor J. C. (Tito) Scaiano on the occasion of his 60th birthday.

Abbreviations: BPA, 9,10-(anthrylenedimethylene) bis (1-piperidine); BDEA, 9,10-bis(N,N-[2-hydroxyethyl]aminomethyl)anthracene; CD, circular dichroism; LD, linear dichroism.



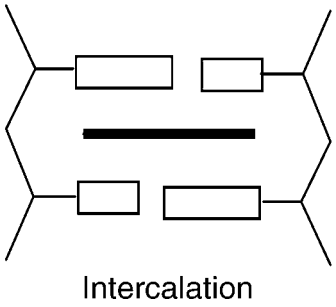


Chart 1. Two major binding modes of small ligands to the DNA double helix. In the case of the anthracene probes, only one geometry for groove binding is shown where the short axis of the anthracene ring system is aligned with the groove walls. The small ovals in the front and the back represent the substituents on the anthracene ring.

The cyclic side chains of BPA are expected to provide some steric blockage for intercalation, but these may not be large enough to completely inhibit intercalation. For example, intercalation of nogalamycin, which has large nonplanar substituents, is well recognized (16,17). The side chains of BDEA, in contrast to BPA, may provide greater steric blockage due to their additional conformational entropy. Steric demand is not the only difference between these substituents. The OH groups of BDEA may participate in H-bonding with appropriate functions on the helix and stabilize specific binding modes. Substituents are placed at 9,10 positions of the anthryl ring such that when the aromatic moiety intercalates these functionalities may occupy the grooves.

The binding properties of BPA/BDEA may also depend on the DNA sequence. The minor groove dimensions, H-bonding locations, polarity, and flexibilities of GC and AT sites differ significantly (18,19). Large, branched substituents on the anthracene ring are suggested to disfavor intercalation at GC sites due to the steric clash with the amino group of guanine in the minor groove (4,20,21). Current results show that substituents of BPA and BDEA are appropriate to evoke considerable differences in their affinities for GC and AT sequences.

Data are presented below to establish a clear distinction between the spectroscopic signatures of the binding modes of two anthracene derivatives. These signatures allowed the tentative assignment of specific features to selected binding modes.

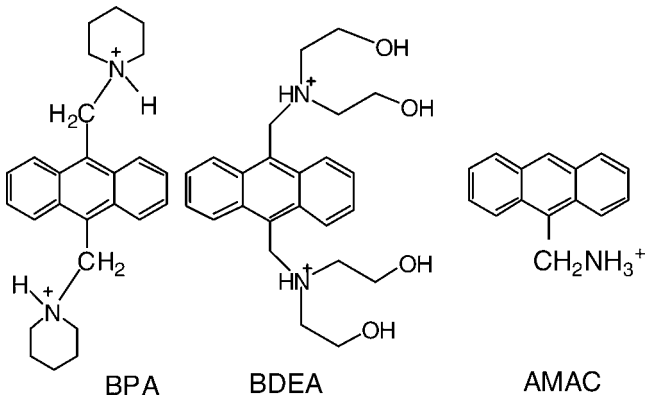


Chart 2. The structures of the anthrylene derivatives investigated in this study.

MATERIALS AND METHODS

DNA samples. DNA samples (CT DNA, poly[dA-dT] and poly[dG]·poly[dC]) were purchased from Sigma Chemical Co. (St. Louis, MO), and CT DNA was purified according to published protocols (22). All solutions were prepared in Tris buffer (5 m*M* Tris, 50 m*M* NaCl, pH 7.2), unless noted otherwise, and all DNA concentrations are expressed in DNA base pairs.

Synthesis of the anthryl derivatives. BPA (9,10-[anthrylenedimethylene]bis[1-piperidine]) was synthesized by a suitable modification of the reported procedure (23). Piperidine (2 equivalents, 5 mmol) was reacted with 9,10-bis(bromomethyl) anthracene (1 g, 2.7 mmol) in chloroform (10 mL), at room temperature. After 48 and 54 h, additional amounts of piperidine (1 equivalent each) were added. The solvent was removed under reduced pressure and the residue washed with several portions of hexane. The product was purified by column chromatography on alumina. Elution with a gradient of methanol/ethyl acetate from 5:95 to 100:0 yielded silky yellow needles (0.5 g): mp 206–208°C (literature value 197–198°C, crystallization from benzene) (24); ^{1}HNMR (CD_3OD) ppm 8.55 (m, 4), 7.77 (m, 4), 5.46 (s, 4), 3.45 (m, 4), 3.35 (m, 4), 1.80 (m, 4), 1.65 (m, 6), 1.52 (m, 2); ^{13}CNMR (CD_3OD) ppm 133.06, 129.10, 126.27, 126.07, 55.54, 53.90, 24.06 and 22.67. Mass spectrum indicated M^+ ion peak at (*m*/*z*) 373.

Synthesis of BDEA (9,10-bis(N,N-[2-hydroxyethyl]aminomethyl)anthracene). Excess of diethanolamine (2 mL) was added to 9,10-bis(bromomethyl) anthracene (0.455 g) in $CHCl_3$ (15 mL). After 24 h, the reaction mixture was dried under vacuum and washed with hexane. The product was purified by column chromatography (20–230 mesh silica gel, elution with a 1:1 mixture of $CHCl_3$:ethanol) to give pure sample (62% yield) mp 172–175°C (reported value 167°C) (25): ^{1}HNMR (CD_3OD, 12 m*M* NaOD) ppm 8.52 (m, 4), 7.41 (m, 4), 4.60 (s, 4), 3.40 (t, 4), 2.67 (t, 4); ^{13}CNMR (CD_3OD) ppm 132.60, 132.19, 126.85, 126.53, 61.12, 57.36 and 52.88. Mass spectrum indicated M+ ion peak at (*m*/*z*) 413.

Both BPA and BDEA samples were converted to the corresponding hydrochlorides by treating with dilute HCl. The first pKa values of similar anthracene–methylamine derivatives are in the range of 6–6.5 (25). Hence, at least one of the two nitrogens of our probes is expected to be protonated in Tris buffer (pH 7.2).

Isothermal titration calorimetry

The binding constants with CT DNA were estimated by isothermal titration method by VP-ITC from MicroCal Inc. (Northampton, MA) at 25°C, as described earlier (26). Predetermined volumes of the probe were injected into the cell and heat released/absorbed has been recorded. The data were analyzed by Origin software (v.5.0, Microcal Inc.) to estimate the binding parameters (binding constant, site size, as well as enthalpy, entropy and free-energy changes).

The heat released or absorbed (Q) during the titration is related to the molar heat of ligand binding (ΔH), volume of the cell (V_o), ligand

concentration (X_t), DNA concentration (M_t), the intrinsic binding constant (K_b) and the number of binding sites (n) by (Eq. 1) (27):

$$Q = 0.5(nM_t\,\Delta HV_o)[1 + (X_t/nM_t) + (1/nK_bM_t) - \sqrt{[(1 + (X_t/nM_t) + (1/nK_bM_t))^2 - 4X_t/nM_t]}] \quad (1)$$

Absorption studies. The absorption spectra were recorded on HP 5750 spectrophotometer by using 1 cm path length cuvettes. The probe solution was titrated by adding 5 μL increments of CT DNA solution, while monitoring the absorbance in 300–500 nm region. Total volume of CT DNA added did not exceed 10% of the sample volume, and the spectra have been corrected for volume changes. The data were analyzed by Scatchard equation (Eq. 2) to estimate the binding constants as described earlier (26). In Eq. 2, C_f is the concentration of the free probe, $r = C_b$/[DNA], C_b is the concentration of bound probe, K_b is the binding constant and n is the binding site size.

$$1/C_f = K_b n(1/r) - K_b \quad (2)$$

Induced circular dichroism studies (CD). The CD spectra were recorded on a Jasco J-710 or J-715 spectropolarimeter interfaced with a Dell personal computer (Jasco Inc., Easton, MD). Solutions containing probe and DNA were placed in a quartz cuvette (1 cm path length for 325–425 nm, and 0.2 cm path length for 200–325 nm) and data were collected with 1 nm bandwidth, 20 millidegrees sensitivity, 4 s response time.

Linear dichroism studies (LD). Flow linear dichroism spectra were recorded on a Jasco J-715 spectropolarimeter adapted for linear dichroism (LD) measurements with the use of a cylindrical flow cell in which the flow is generated by the rotation of an inner quartz cylinder, spaced 0.5 mm from a stationary steel/quartz cylinder (28). The ratio of the concentration of DNA to that of the probe was varied by the addition of a concentrated solution of the probe as well as an appropriate volume of DNA solution to the cell. Thus, the DNA concentration and the sample viscosity are kept constant throughout the titration. All spectra were recorded at pH 6.9 (1 m*M* sodium cacodylate, 50 m*M* NaCl). Absorbance spectra of identically prepared samples were collected on a Jasco V-550 spectrometer in a 1 cm path length cuvette to calculate LD^r spectra. The equation $LD^r = LD/A = 3/2S(3\cos^2\alpha - 1)$ was used, where A is the isotropic absorbance, S is the orientation parameter (0 for random and 1 for perfect orientation), and α is the angle between the transition moment polarization (which is long axis at 257 nm and short axis at 380 nm for anthracene chromophore) and the DNA helix axis.

DNA melting experiments. Helix melting studies were carried out by a Hewlett-Packard 8452A diode array spectrophotometer as described earlier (26). The T_m values are obtained as the midpoints of the thermal transitions observed at 260 nm.

Differential scanning calorimetry (DSC). The DSC thermograms were recorded on a Model 6100 Nano II differential scanning calorimeter (Calorimetry Sciences Corp., Lindon, UT), as described earlier (29). The excess molar heat capacity was calculated from the DSC data, and the thermodynamic parameters estimated. Each value represents average of at least three measurements.

RESULTS AND DISCUSSION

The binding properties of BPA and BDEA (Chart 2) are investigated by calorimetric, spectroscopic and helix melting studies. Binding mode preferences and selectivities are evaluated with the use of CT DNA, poly(dG)-poly(dC) and poly(dA-dT). Current data point to clear differences in the DNA binding properties of BPA and BDEA. Tentative assignment of specific spectral features to particular binding modes is based on the following observations.

Isothermal titration calorimetry. The intrinsic binding affinity and the corresponding thermodynamic parameters were measured directly by isothermal titration calorimetry (Fig. 1A). Addition of small volumes of a concentrated solution of BPA (264 μ*M* 25°C) to a solution of CT DNA (100 μ*M*) resulted in heat release. The binding saturated at BPA:DNA concentration ratio of ~0.2. In order to account for the contribution of the heat of dilution to the data, a separate titration was carried out in which a solution of BPA (264 μ*M*) was added to buffer (Fig. 1A, top curve). Similarly, the heat of dilution of CT DNA was also measured (data not shown). The area under each peak was integrated, and the heats of dilution subtracted. These data are shown in Fig. 1B.

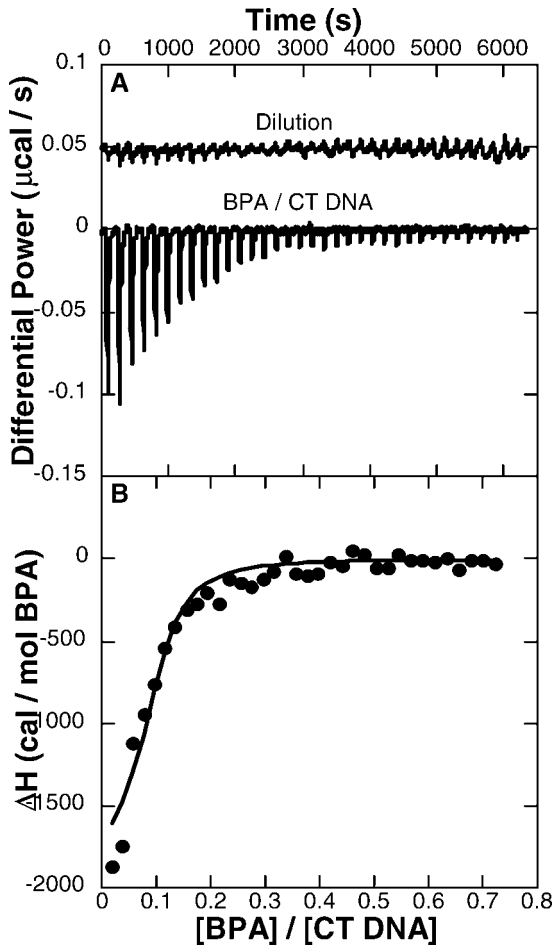


Figure 1. (A) The isothermal titration calorimetric data on the binding of BPA (264 μ*M*) to CT DNA (100 μM) at 25°C, Tris buffer (50 m*M* NaCl, pH 7.2, top panel, bottom curve). Similar titration of BPA solution into the buffer resulted in weak heat release (top curve, top panel). (B) The areas under each peak were integrated, and the heat release was corrected for dilution (bottom panel). The solid line corresponds to the best fit to the data.

The ITC data are fitted to a single binding site model (Eq. 1) (30) and the best fits to the data are indicated: $K_b = 6.5 \pm 1.1 \times 10^5\ M^{-1}$; $\Delta H = -1.92 \pm 0.04$ kcal/mol; $\Delta S = 20.1 \pm 0.2$ cal/molK; and site size of ~9 base pairs. The data were analyzed with several different initial guesses, and the resulting fits gave consistent values. Note that K_b of BPA is nearly an order of magnitude larger than that of AMAC measured at similar ionic strength, and pH (26). Exothermic binding of other anthracene derivatives to CT DNA was also reported earlier (26,29).

In contrast to the exothermic binding of BPA, titration of BDEA (200 μ*M*) with CT DNA (200 μ*M*) did not indicate any heat release (at 25°C or at 40°C, data not shown). Note that the binding of AMAC to CT DNA was also thermoneutral. Therefore, spontaneous binding of these probes is entropically driven ($\Delta S > 0$) and part of the entropy increase is likely due to the release of water molecules from the molecular surfaces that are brought together during the interaction (31).

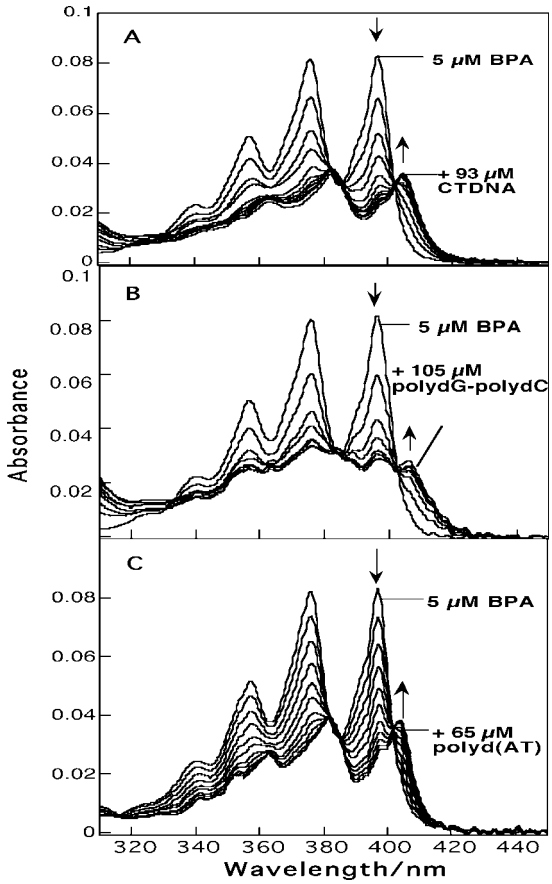


Figure 2. Absorption spectra of BPA (5 μ*M*) recorded in the presence of increasing concentrations of (A) CT DNA (0–93 μ*M*); (B) poly(dG)-poly(dC) (0–105 μ*M*); and (C) poly(dA-dT) (0–65 μ*M*). Clear redshifts in the absorption spectra, in the presence of excess of DNA, are clearly demonstrated.

Absorption studies. The absorption spectra of BPA (5 μ*M*) were recorded in the presence of increasing concentrations of CT DNA (0–93 μ*M*, Fig. 2A). The spectra were corrected for small changes in volume during the titration. The data indicated strong hypochromism (67% at 397 nm), and a substantial redshift (498 cm^{-1}, Fig. 2A). Note that the peak at 397 nm is nearly completely abolished at the higher DNA concentrations, and the vibronic bands of the bound form are well resolved. The full width at half maximum (FWHM) of the 0–0′ band is increased from 775 to 1070 cm^{-1} (S1). The presence of isosbestic points at 402, 386 and 382 nm* indicate that there are at least two spectroscopically distinct chromophores (free and bound). The spectral features observed here are generally assigned to intercalation (mode I) (9,11), and this assignment is further discussed below.

Titration of BPA (5 μ*M*) with poly(dG)·poly(dC) (0–105 μ*M*) resulted in absorption changes similar to those observed for CT DNA (Fig. 2B), but the behavior is not identical. Although the absorption at 397 nm decreased with increasing DNA concentration, this peak persisted even at high DNA concentrations (105 μ*M*). The vibronic bands are not well resolved, but remained broad. However, strong hypochromism (58% at 397 nm), bathochromic shift (436 cm^{-1}), and isosbestic points at 402, 386 and 382 nm are noted. In addition to the redshifted, hypochromic, broad spectral component (mode I) there is another component contributing to these spectra.

Addition of more DNA did not resolve the vibronic bands, and spectral features similar to these are also reported with AMAC (9). Another binding mode (mode II) with peak positions that are nearly the same as those of the free chromophore becomes important.

Titration of BPA with poly(dA-dT) indicated spectral changes that are similar to those observed with CT DNA (Fig. 2C). Bathochromism (375 cm^{-1}), hypochromism (62% at 397 nm), isosbestic points (402, 386 and 382 nm) and well-resolved vibronic bands (FWHM, 0–0′ band, 930 cm^{-1}) are consistent with binding mode I. Minor, unshifted peaks persisted at high DNA concentrations, but mode I dominated. It is likely that two spectroscopically distinct binding modes (I and II) exist for BPA, and the relative contributions of these depend on the DNA sequence.

The absorption spectra of BDEA (7, 6, 5 μ*M*) recorded in the presence of DNA polymers indicated interesting changes (Fig. 3A–C). Moderate hypochromism (30–38%), but no significant bathochromism, is indicated with all the three DNA polymers. Multiple isosbestic points are observed with BDEA, in specific cases.† Continuous decreases in the peak absorbances were noted. The 0–0′ band was not broadened to any significant extent, upon binding to any of the three DNA polymers (S1), and there is no clear indication of a redshifted spectrum. The unshifted, hypochromic spectrum of BDEA cannot be attributed to the free probe, and these features are consistent with mode II, discussed above. Therefore, the spectral features of BDEA/DNA are clearly distinct from those of BPA/DNA.

Binding isotherms. The absorption data were used to construct binding plots (Fig. 4), and the plot was analyzed with the use of Eq. 2. Linear fits to the BPA data resulted in the corresponding binding constants: 1.6×10^6, 1.2×10^6, and $2.3 \times 10^6 M^{-1}$, with CT DNA, poly(dG)·poly(dC) and poly(dA-dT), respectively (Table 2). Note that the affinity for poly(dG)·poly(dC) is similar to that of CT DNA, and BPA prefers binding to poly(dA-dT) by only a factor of 2. These affinities are much higher than those reported for monosubstituted anthracene derivatives (4), but smaller than that of a triply charged derivative (7).

The binding constants of BDEA with CT DNA and poly(dG)·poly(dC) are nearly equal (2.4×10^6 M^{-1} to 2.8×10^6 M^{-1}, S2, Table 2) but greater than those of BPA. The acyclic side chains of BDEA did not reduce its affinity for these DNAs. The binding constant of BDEA with poly(dA-dT) is nearly 10 times less than with poly(dG)·poly(C), which indicates a high selectivity.

Circular dichroism studies. Binding of achiral chromophores to the DNA double helix often results in induced circular dichroism (ICD), and this is due to the asymmetric environment of the probe at the helix. The intensity and sign of the ICD signal can be used to assign the relative orientations of the anthryl chromophore with respect to the helix axis (32,33). The electronic transition of the anthracene chromophore assigned to absorption in the 320–400 nm

* The spectra intersect over a region of 0.002 absorbance units and on a full scale of 0.1 absorbance units this deviation is ±1%. If there is a third species, its contribution is ~1% or less.

† The BDEA absorption spectra recorded in the presence of CT DNA and poly(dG)·poly(dC) (Fig. 3A,B) indicated isosbestic points at 406 and 404 nm, respectively. In Fig. 3B only one spectrum did not pass through the intersection. In Fig. 3C the spectra passed within a region of 0.002 absorption units on a full scale of 0.05 absorbance units (4% maximum deviation).

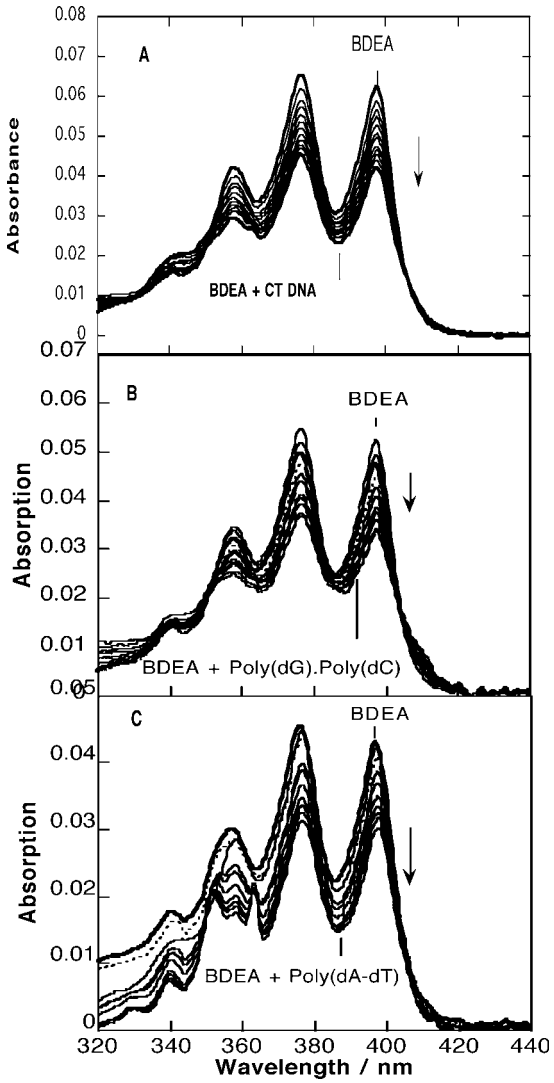


Figure 3. Absorption spectra of BDEA (5 μ*M*) recorded in the presence of increasing concentrations of (A) CT DNA (0–173 μ*M*); (B) poly(dG)-poly(dC) (0, 11, 22, 33, 43, 54, 64, 75, 96 and 105 μ*M*); and (C) poly(dA-dT) (0, 12, 24, 36, 48, 59, 70, 81, 92 and 113 μ*M*). The lack of redshifts in the spectra, in the presence of DNA, is clearly demonstrated.

region is short-axis polarized, and positive ICD spectra in this region are attributed to the intercalative binding of the anthracene chromophore where anthryl short-axis is aligned parallel to the base pair pseudodyad axis. This orientation and ICD signatures were also noted for other anthracene derivatives (34,35). When BPA binds to CT DNA or poly(dA-dT) strong positive ICD bands are noted (Fig. 5). The peak positions of the ICD spectra match with the redshifted absorption peak positions of the bound chromophore, and hence, the positive ICD spectra correspond to binding mode I.

Note that much weaker ICD is indicated with BPA/poly(dG)-poly(dC). Most likely, this reduction in intensity is due to the contributions of mixed binding modes (I and II), where one of the modes has very weak signals or no signals at all, or signals of the opposing sign. Note the differences between the peak positions of the ICD and the absorption spectra with this polymer.

In contrast to BPA, the ICD spectra of BDEA recorded in the presence of CT DNA and poly(dA-dT) indicated very weak

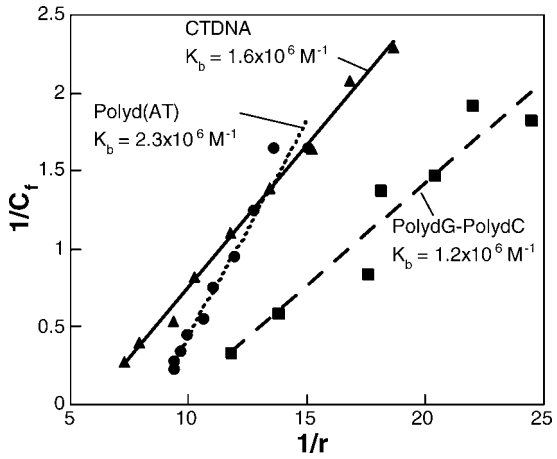


Figure 4. Scatchard plots for the binding of BPA to CT DNA (triangles), poly(dG)-poly(dC) (squares) and poly(dA-dT) (circles).

signals (by a factor of 6, Fig. 6) and ICD spectra of BDEA/poly(dG)·poly(dC) were even weaker, or nonexistent. The peak positions of the BDEA/CT DNA ICD spectra are at wavelengths (405, 385 and 365 nm) that do not match with the corresponding absorption peaks (397, 376 and 357 nm). This discrepancy and the weaker signals clearly suggest that the ICD spectra are not due to the major binding mode (II) but due to another, minor mode. Similarly, the weak ICD spectra of BDEA/poly(dA-dT) also arise from a minor binding mode. The ICD of BDEA clearly show that mode II (no ICD) is distinct from mode I (strong ICD).

The DNA CD at 280 nm, with and without the probe, was useful in learning about the helix unwinding by the probe. For example, the UV CD of CT DNA (300 μ*M*) decreased dramatically by the addition of BPA (75 μ*M*) but not by BDEA (75 μ*M*) (S4). This loss of DNA CD was suggested to be characteristic of helix unwinding (10,36). Thus, mode I is associated with significant unwinding of the helix, whereas mode II is not.

Linear dichroism (LD) studies. Flow linear dichroism data are useful for the determination of the orientation of the chromophore with respect to the helix axis, and this has been used to help characterize the binding modes. BPA indicated strong negative LD peaks (300–450 nm) in the presence of CT DNA or poly(dA-dT) (Fig. 7). The intensities were much weaker with poly(dG)-poly(dC). The weaker intensity in this case is not due to a significantly weaker binding (Table 2), nor to DNA bending (unscaled LD spectra at 280 nm where there is no ligand absorbance were

Table 1. Absorption spectral properties of BPA and BDEA bound to DNA. The extinction coefficients of BPA and BDEA at 397 nm (0–0′ band) are 16400 and 8600 M^{-1} cm^{-1}, respectively

Probe	DNA	λ_{bound} (nm)	ΔE (cm^{-1})	ε_{bound} (M^{-1} cm^{-1})*	% Hypochromism (397 nm)
BPA	CT DNA	405	498	4800	67
	Poly(dG)-Poly(dC)	404	436	5800	58
	Poly(dA-dT)	403	375	5100	62
BDEA	CT DNA	397	0	6000	30
	Poly(d G)-Poly(dC)	397	0	5700	34
	Poly(dA-dT)	397	0	6200	28

* Errors in the measurements are less than ±5%.

Table 2. The DNA binding constants and quenching constants for BPA and BDEA with three different DNA polymers

Probe	DNA	K_b ($10^6 M^{-1}$)
BPA	CT DNA	1.6
	Poly(dG)·poly(dC)	1.2
	Poly(dA-dT)	2.3
BDEA	CT DNA	2.4
	Poly(dG)·poly(dC)	2.8
	Poly(dA-dT)	0.3

10–20% lower than that with the ligands), so it must be related to the orientation of the anthracene on the DNA. The LD in the 300–450 nm region is assigned to the polarization along the short axis of the anthryl probe, and the negative LD indicates that the short axis of BPA is more perpendicular to the helix axis than it is parallel. This characteristic has often been attributed to the intercalative mode (37). Groove binding where the short axis of the probe is nearly perpendicular to the helix axis (Chart 1) is another possible configuration, which can result in negative LD in the 380 nm region. The absence of a strong LD spectrum with poly(dG)-poly(dC) indicates that there are additional possibilities (multiple modes) for the orientation of the probe.

Groove binding and intercalation, in this case, may therefore result in negative LD. With CT DNA, a second binding mode becomes evident at 9:1 or greater DNA:BPA ratio. The ligand stiffens the DNA at lower drug load than this, although not dramatically. Reduced LD, determined by $LD_r = LD/A_{iso}$, was calculated for the samples where it seemed only one mode was operative. The ligand band at 380 nm is redshifted and the ligand is oriented more perpendicular to the helix axis than the average base pair; this is consistent with the intercalator locally stiffening the DNA (data not shown). With poly(dA-dT), the same redshift as with CT DNA was observed, and the LD does not indicate a second binding mode, even at low drug loads.

In the DNA absorption regions (<300 nm) strong negative bands are seen with all DNAs (poly[dG]·poly[dC] is less well oriented due to its shorter length). The LD around 260 nm is from

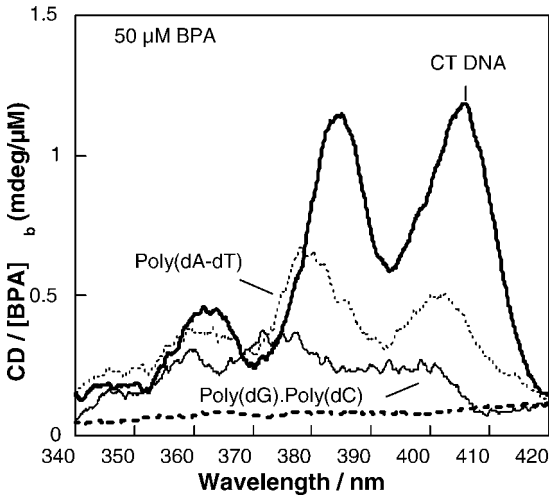


Figure 5. Circular dichroism spectra of BPA (50 μ*M*, no DNA, dashed line), recorded in the presence of CT DNA (400 μ*M*, boldface line), poly(dG)-poly(dC) (400 μ*M*, lightface line) and poly(dA-dT) (400 μ*M*, dotted line).

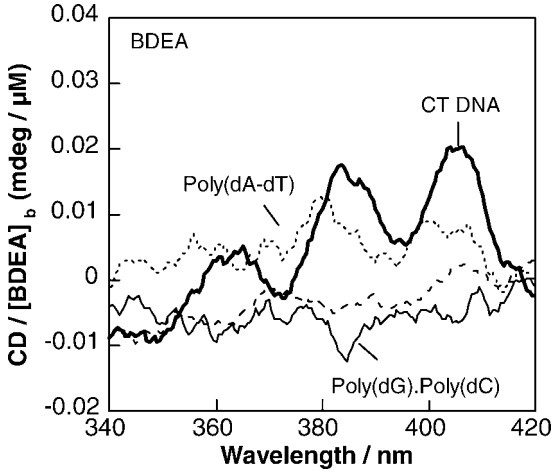


Figure 6. Circular dichroism spectra of BDEA recorded in the presence of CT DNA (420 μ*M*, boldface line, 28 μ*M* BDEA), poly(dG)·poly(dC) (214 μ*M*, lightface line, 20 μ*M* BDEA) and poly(dA-dT) (318 μ*M*, dotted line, 20 μ*M* BDEA). The dashed line is the baseline (20 μ*M* BDEA) recorded in the absence of DNA.

the polarization of the transition of the DNA bases and the long axis of the anthryl probe (254 nm). For an intercalator, the LD magnitude in the 260 nm region should increase with increasing drug load because of the stiffening of the DNA helix by the intercalator (38). The difference in the increase of LD in this region when BPA binds to DNA may be correlated with the degree of intercalation. The magnitude of LD is very small with poly(dG)·poly(dC), and the broad LD bands are indicative of weak intercalation. However, the 258 nm dip in LD is quite pronounced even at DNA:BPA ratio of 24:1, indicating that a second nonintercalative mode is operating even at low drug loads with this polymer. The evidence suggests weak intercalation is the first mode but that the second binding mode, possibly groove binding, dominates quickly. The increase in LD magnitude is more pronounced in poly(dA-dT) and CT DNA than in poly(dG)·poly(dC). This is consistent with predominance of intercalation of BPA into poly (dA-dT) and CT DNA than into poly(dG)·poly(dC).

BDEA LD spectra are shown in Fig. 8 after normalization at 280 nm. The signals were only 1–2% higher than that of the free DNA in contrast to the situation for BPA, where the DNA was significantly stiffened. For each DNA the short axis polarized 380 nm band has a negative LD signal, except at very high loading (3.5:1) with poly(dA-dT). The 257 nm region also has a net negative ligand contribution. With CT DNA only, there is clear evidence of a positive contribution at 257 nm at higher loading. All in all, the LD confirms that BDEA binds to DNA and gives no clear evidence in support of intercalative binding, and groove binding dominates (Chart 1). Both long and short axis polarized transitions have negative signals, which is consistent with a mode where the anthracene short axis is nearly perpendicular to the helix axis. A favorable interaction of the H bonding and cationic functions with the phosphate backbone can stabilize this unusual configuration. Thus, LD data indicate intercalation of BPA into CT DNA and poly(dA-dT) and no intercalation of BDEA.

Helix melting studies. Binding of the anthracene probes to the DNA helix may improve the helix stability, and intercalation of the anthracene probes indicated increased stability (4,9,29). In the case of BPA, the T_m of CT DNA was increased from 73 to >95°C (Fig. 9A), and similar increase was also observed with

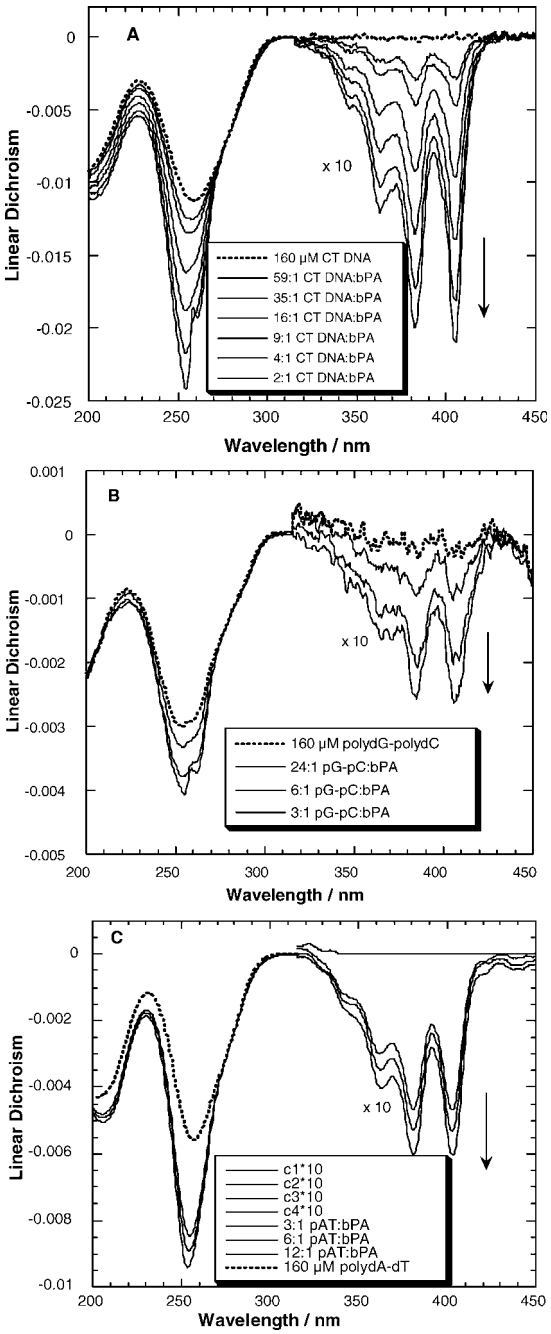


Figure 7. Linear dichroism spectra of BPA (normalized at 280 nm) recorded in the presence of (A) CT DNA (160 µ*M*); (B) poly(dG-dC) (160 µ*M*); and (C) poly(dA-dT) (160 µ*M*). A solution of BPA was titrated into the DNA solution, which resulted in different ratios of DNA to BPA concentrations, and additional DNA was added to keep its concentration constant throughout the titration.

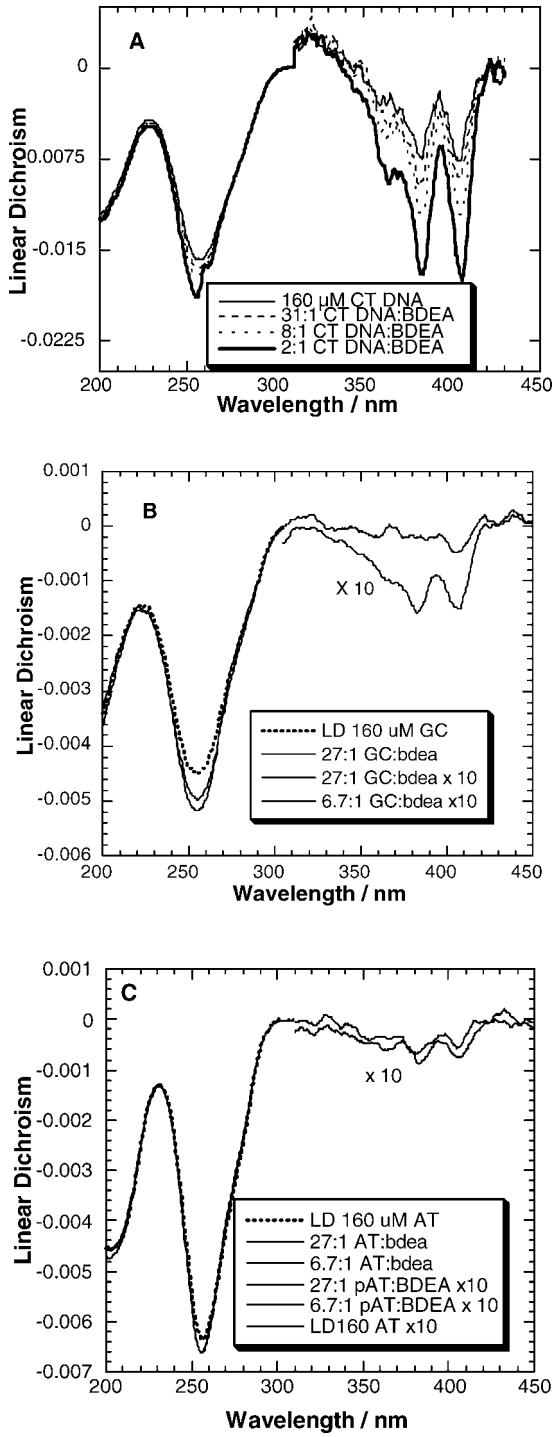


Figure 8. Linear dichroism spectra (normalized at 280 nm) of BDEA recorded in the presence of (A) CT DNA (160 µ*M*) (B) poly(dG-dC)(160 µ*M*) and (C) poly(dA-dT) (160 µ*M*). A solution of BDEA was titrated with the DNA solution, which resulted in different ratios of DNA to BDEA concentrations, and additional amounts of DNA were added to keep DNA concentration constant throughout the titration.

poly(dG)·poly(dC) (Fig. 9B). Note that BPA increases the T_m of poly(dA-dT) from 47 to >80°C (Fig. 9C). By far, BPA is one of the best anthracene derivatives, and showed T_m increases >30°C (Table 3). Large T_m increases are also noted with other anthracene derivatives, and these were assigned to binding mode I (4,9,26,29).

In contrast to these data, no increases in T_m are noted with BDEA (Fig. 10). For example, the helix melting curve of CT DNA (60 µ*M*) recorded in the presence of BDEA (15 µ*M*, Tris buffer) indicated no change in T_m (Fig. 10A). This is extremely unusual for an anthracene derivative. Similar unchanged T_m values are also noted with poly(dG)·poly(dC) as well as poly(dA-dT) (Fig. 10B,C). These data clearly distinguish between the properties of BPA and BDEA. Although mode I stabilizes the helix, mode II had no influence.

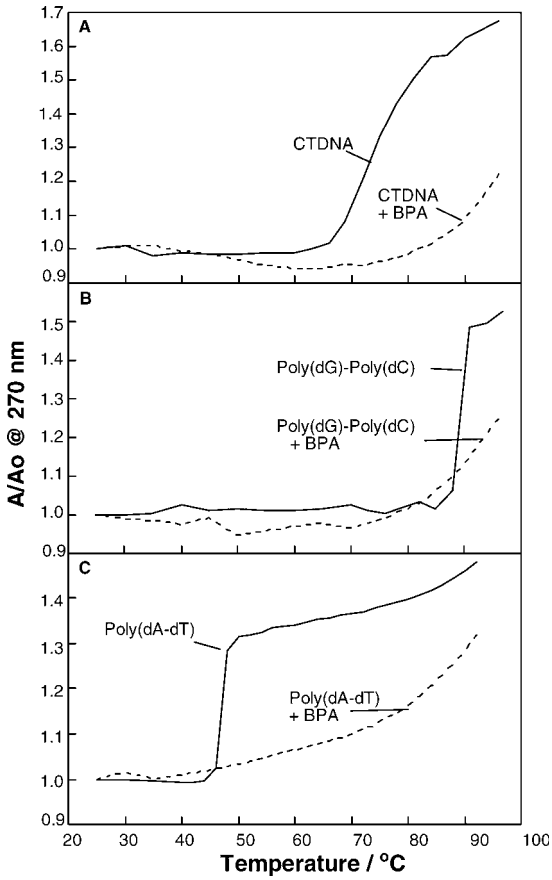


Figure 9. Helix melting curves of BPA (30 μ*M*)/DNA (30 μ*M*). (A) CT DNA; (B) poly(dG-dC); and (C) poly(dA-dT), all recorded at 10 m*M* NaCl (5 m*M* Tris buffer). The T_m is increased in all three cases.

DSC studies. The DSC thermogram of CT DNA (60 μ*M*) indicated a strong transition at 79 ± 0.6°C (Fig. 11) and several small satellite peaks which are attributed to short, repetitive stretches of GC sequences (39). The integrated area under the thermogram is the enthalpy of denaturation, and the thermodynamic parameters obtained (Table 4) agree with reported values (40). The thermogram of a mixture of BPA (10 μ*M*) and CT DNA (60 μ*M*) indicated a shift to higher temperatures (T_m is 86 ± 0.6°C, Fig. 11). Note that the high-temperature satellite peaks of CT DNA are missing, and that one such satellite peak appeared down at ~70°C. It is interesting to note that certain small stretches of CT DNA was destabilized by BPA to a large extent.

The enthalpy of helix denaturation decreased from 8.9 ± 0.2 kcal/mol for CT DNA to 5.9 ± 1.0 kcal/mol (Table 4) for BPA/CT DNA. Such decreases in denaturation enthalpies are reported earlier (41). The DSC data clearly show that binding mode I is

Table 3. The sign of circular dichroism spectra (320–450 nm) and ΔT_m values for BPA and BDEA in the presence of three different DNA polymers

	CD			ΔT_m (°C)		
Probe	CT DNA	Poly(dG)·poly(dC)	Poly (dA-dT)	CT DNA	Poly(dG)·poly(dC)	Poly (dA-dT)
AMAC	+ve	+ve	+ve	7	>10	22
BPA	+ve	Weak	+ve	>20	>10	>30
BDEA	Weak	No CD	Very weak	0	0	0

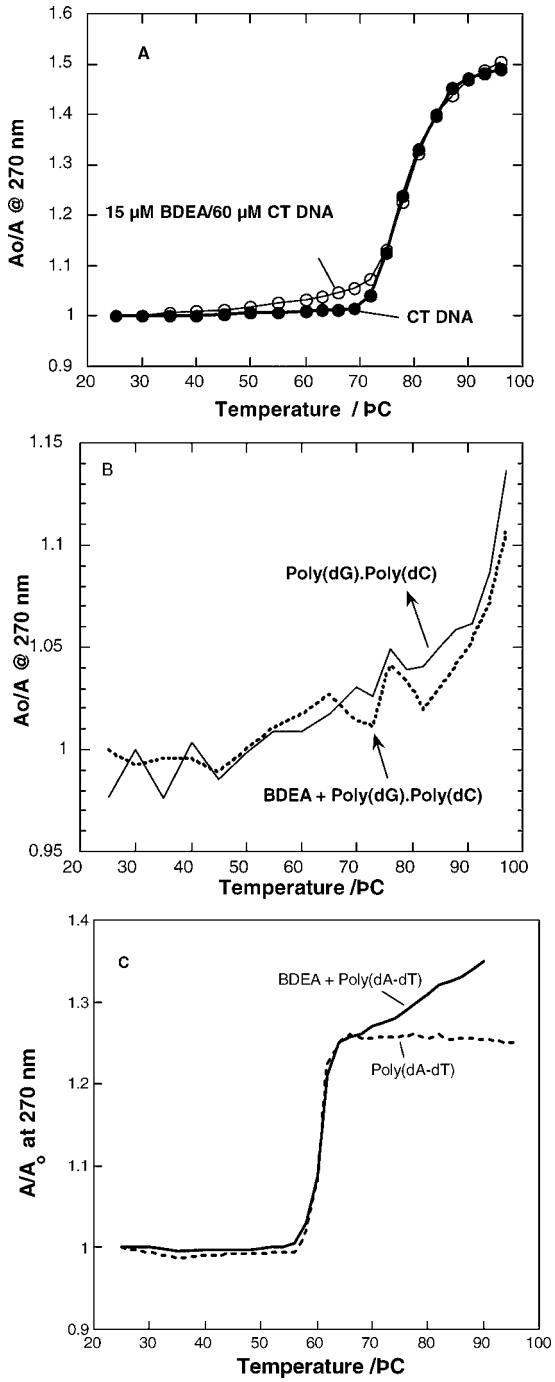


Figure 10. The helix melting curves for (A) CT DNA (60 μ*M*); (B) poly(dG)·poly(dC) (40 μ*M*); and (C) poly(dA-dT) (100 μ*M*) recorded in the presence and absence of BDEA (15 μ*M*, Tris buffer, 50 m*M* NaCl pH 7.2). The curves with and without BDEA have nearly the same T_m values.

associated with increased helix stability, and similar data with BDEA indicated no change in the helix stability, T_m or the enthalpy of denaturation. Therefore, we conclude that BDEA does not stabilize the helix, which is consistent with the helix melting studies.

Low-ionic-strength studies. To further evaluate if BDEA is capable of binding by mode I, we tested its binding properties at lower ionic strength (10 m*M* NaCl, 1 m*M* Tris). This is because at lower ionic strengths the DNA helix expands owing to increased

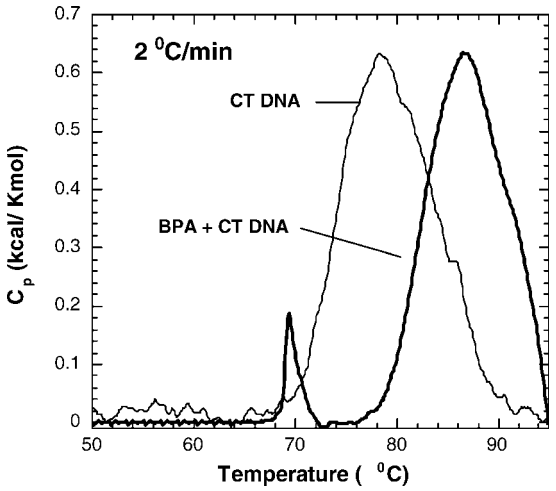


Figure 11. The DSC thermograms of CT DNA (60 μ*M*) and a mixture of BPA (10 μ*M*) and CT DNA (60 μ*M*) (5 m*M* Tris, 50 m*M* NaCl, pH 7.2).

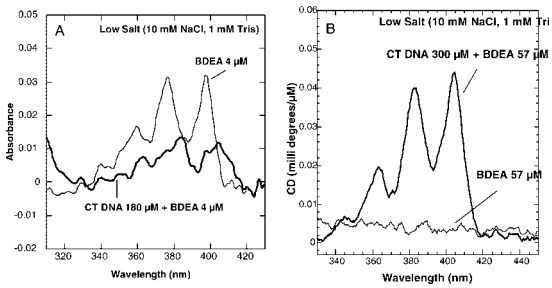


Figure 12. (A) Absorption spectra of BDEA (4 μ*M*) recorded in the presence of CT DNA (180 μ*M*, 10 m*M* NaCl, 1 m*M* Tris pH 7.2); (B) the ICD spectra of BDEA (57 μ*M*) recorded in the presence of CT DNA (300 μ*M*, 10 m*M* NaCl, 1 m*M* Tris pH 7.2).

phosphate–phosphate repulsions (42), and this may alleviate some steric clash with the ligand. In addition, the decreased ionic strength is expected to improve the electrostatic interactions with the helix. Addition of CT DNA (180 μ*M*, 10 m*M* NaCl, 1 m*M* Tris) to BDEA (4 μ*M*, 10 m*M* NaCl, 1 m*M* Tris) resulted in strong hypochromism (66% at 404 nm), and a redshift (436 cm^{-1}) in the absorption spectrum (Fig. 12A). These changes are characteristic of mode I, and they are remarkably similar to those observed with BPA/CT DNA (Fig. 2A, 50 m*M* NaCl). Therefore, BDEA is capable of binding via mode I.

Binding of BDEA to CT DNA at lower ionic strengths was also examined in CD studies to further clarify the binding mode. As anticipated, strong ICD bands are noted (Fig. 12B), when CT DNA (300 μ*M*) was added to BDEA (57 μ*M*). Note that the ICD peak positions correspond to the redshifted absorption peak positions of the bound form (Fig. 12A).

These data clearly show that BDEA is capable of interacting with CT DNA in a manner similar to that of BPA, with essentially the same spectral signatures. But this occurs at a lower ionic strength, and ionic strength has a strong influence on the preferred binding mode of BDEA.

CONCLUSIONS

Several lines of investigations draw clear distinctions between the binding properties of the two probes. Two major binding modes are identified. The greater conformational entropy of the side chains of BDEA, in comparison to BPA, make it sterically slightly more demanding. The H-bonding groups on this ligand are potential sites for favorable interaction with the helix. These two probes are similar in terms of the charge, geometry and shape. The above data clearly show that there are at least two distinct binding modes. For the first time, the differences between the two modes are clearly established and their spectroscopic signatures identified. The characteristics of mode I are more consistent with intercalation than groove binding, whereas mode II is more likely to be groove binding than intercalation.

These assignments are supported by the following characteristics. Mode I indicated: (1) large redshifts in the absorption spectra (375–500 cm^{-1}); (2) extensive hypochromism (>55%); (3) substantial broadening of the vibronic bands (150–300 cm^{-1}); (4) strong positive ICD signals in the 320–420 nm region; (5) strong decrease in the CD DNA at 280 nm; (6) strong negative LD in the 320–420 nm region; (7) large increases in T_m ($\Delta T_m \sim$10–35°C); and (8) large decrease in the helix denaturation enthalpy. The majority of a number of different anthracene probes, with small substituents, indicated similar changes in their properties, and these have been assigned to the intercalative binding mode (4,6,7,9–12,23,26,29).

Intercalation of the anthracene ring would allow for strong stacking interactions with the DNA base pairs. The low polarizability of the anthracene ring system is expected to strengthen such interactions. The close proximity of the π electrons to the base pairs would result in excitonic splitting of the energy levels, redshifts, and considerable broadening of the vibronic bands. Numerous properties reported here for mode I are indicative of intercalation of BPA.

Current data clearly identify another binding mode for the anthracene derivatives, for the first time, which indicated distinct properties. Little or no redshift in the absorption spectra, smaller hyperchromism (×38%), no significant changes in the FWHM of the vibronic bands, substantially weaker or no CD, weak LD signals, and no changes in T_m or the enthalpy of helix denaturation were observed. We tentatively assign this mode to be groove binding. Similar properties are also noted with another anthracene derivative containing a phenyl substituent (11).

At low salt conditions, the binding of BDEA to CT DNA clearly indicated spectral properties that are characteristic of mode I. These data clearly show that BDEA is capable of binding via mode I but prefers mode II at normal salt conditions. The simple correlation between the binding mode and the absorption spectral changes, identified here, facilitates the analysis of the binding modes of

Table 4. Thermodynamic parameters for the denaturation of CT DNA (60 μ*M*) in the presence of BPA (10 μ*M*)

Probe/DNA	ΔH (kcal/mol)	Δ*S* (kcal/kmol)	T_m (°C)
CT DNA	8.9 ± 0.2	0.025 ± 0.002	79 ± 0.6
BPA/CT DNA	6 ± 1	0.016 ± 0.002	86 ± 0.6

Table 5. The dominant DNA binding modes for BPA and BDEA

Binding mode	CT DNA	Poly(dG)·Poly(dC)	Poly(dA-dT)
Intercalation	BPA	BPA	BPA
Groove binding	BDEA	BDEA/BPA	BDEA

other anthracene derivatives. Here is the first example of an anthracene derivative that showed a strong preference for groove binding. Preliminary data from our laboratory, with several sterically crowded anthracene derivatives, indicate that these also exhibit such binding properties.‡ Such tuning of the preferred binding mode by structure or ionic strength is attractive.

Groove binding of BDEA (Chart 1) would position the H-bonding substituents in proximity for interaction with the phosphate groups, which may stabilize this unusual mode. It is also possible that the anthracene derivative may cover the groove with its short axis, or the long axis, bridging the minor groove, and the two orientations cannot be distinguished at this time. Another issue regards the binding in the minor groove versus major groove, which will be investigated in future studies.

The DNA sequence dependence data further strengthen our assignments. For BPA, the contributions of mode I to the overall binding is substantially reduced with GC sequences, and mode II begins to become significant. The wider minor groove of GC sequences can accommodate the branched side chains more readily than AT sequences. BPA indicated the importance of mode I with these sequences.

The side chains of BDEA promote the strong discrimination between AT and GC sites, and this resulted in a 10-fold selectivity for GC sites. Note that this is not due to enhanced affinity for GC sites, but due to a decrease in affinity for the AT sites. This selectivity is due to the high preference of BDEA for groove binding.

Acknowledgements—Financial support from the National Science Foundation (DMR-0300631) is gratefully acknowledged.

REFERENCES

1. Nielsen, P. E. (1991) Sequence-selective DNA recognition by synthetic ligands. *Bioconjug. Chem.* **2**, 1–12.
2. Carter, S. K., D. Rall, P. Schein, R. D. Davis, H. B. Wood, R. Engle, J. P. Davignon, J. M. Venditti, S. A. Schepartz, B. R. Murray and C. G. Zubrod (1968) *Pseudourea*. National Cancer Chemotherapy Institute Clinical Brochure No. NSC56054, p. 2.
3. Penta, J. S., M. Rozencweig, A. M. Guarino and F. Muggia (1979) Mouse and large-animal toxicology studies of twelve antitumor agents: relevance to starting dose for phase I clinical trials. *Cancer Chemother. Pharmacol.* **3**, 97–101.
4. Kumar, C.V., E. H. Punzalan, and W. B. Tan (2000) Adenine–thymine base pair recognition by an anthryl probe from the DNA minor groove. *Tetrahedron* **56**, 7027–7040.
5. S. Neidle (2001) DNA minor groove recognition by small molecules. *Nat. Prod. Rep.* **18**, 291–309.
6. Kumar, C. V., W. B. Tan and P. W. Betts (1997) Hexaminecobalt(III) chloride assisted, visible light induced, sequence dependent cleavage of DNA. *J. Inorg. Biochem.* **68**, 177–181.
7. Rodger, A., I. S. Blagbrough, G. Adlam and M. L. Carpenter (1994) DNA binding of a spermine derivative: spectroscopic study of anthracene-9-carbonyl-N1-spermine with poly[d(G-C)·d(G-C)] and poly[d(A-T)·d(A-T)]. *Biopolymers* **34**, 1583–1593.
8. Adlam, G., I. S. Blagbrough, S. Taylor, H. C. Latham, I. S. Haworth, and A. Rodger (1994) Multiple binding modes with DNA of anthracene-9-carbonyl-N1-spermine probed by LD, CD, normal absorption, and molecular modelling compared with those of spermidine and spermine. *Bioorg. Med. Chem. Lett.* **4**, 2435–2440.
9. Kumar, C. V. and E. H. Asuncion (1993) DNA binding studies and site selective fluorescence sensitization of an anthryl probe. *J. Am. Chem. Soc.* **115**, 8547–8553.
10. Rodger, A., S. Taylor, G. Adlam, I. S. Blagbrough and I. S. Haworth (1995) Multiple DNA binding modes of anthracene-9-carbonyl-N1-spermine. *Bioorg. Med. Chem.* **3**, 861–872.
11. Becker, H. C. and B. Norden (1999) DNA binding mode and sequence specificity of piperazinylcarbonyloxyethyl derivatives of anthracene and pyrene. *J. Am. Chem. Soc.* **121**, 11947–11952.
12. Wilson, W. D., Y. H. Wang, S. Kusuma, S. Chandrasekaran and D. W. Boykin (1986) The effect of intercalator structure on binding strength and base-pair specificity in DNA interactions. *Biophys. Chem.* **24**, 101–109.
13. Atkins, P. and J. de Paula (2002) *Physical Chemistry*. Freeman, New York.
14. Joseph, J., E. Kuruvilla, A. T. Achuthan, D. Ramaiah, and G. B. Schuster (2004) Tuning of intercalation and electron-transfer processes between DNA and acridinium derivatives through steric effects. *Bioconj. Chem.* **15**, 1230–1235.
15. Miller, K. J., J. Macrea, and J. F. Pycior (1980) Interactions of molecules with nucleic acids. III. Steric and electrostatic energy contours for the principal intercalation sites, prerequisites for binding, and the exclusion of essential metabolites from intercalation. *Biopolymers* **19**, 2067–2089.
16. Arora, S. K. (1983) Molecular structure, absolute stereochemistry, and interactions of nogalamycin, a DNA-binding anthracycline antitumor antibiotic. *J. Am. Chem. Soc.* **105**, 1328–1332.
17. Neidle, S., L. H. Pearl, and J. V. Skelly (1987) DNA structure and perturbation by drug binding. *Biochem. J.* **243**, 1–13.
18. D. Kahne (1995) Strategies for the design of minor groove binders: a re-evaluation based on the emergence of site-selective carbohydrate binders. *Chem. Biol.* **2**, 7–12.
19. Dervan, P. B. and R. W. Burli (1999) Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **3**, 688–693.
20. Becker, H. C. and B. Nordén, B. (2000) DNA binding thermodynamics and sequence specificity of chiral piperazinecarbonyloxyalkyl derivatives of anthracene and pyrene. *J. Am. Chem. Soc.* **122**, 8344–8349.
21. Sissi, C., L. Bolgan, S. Moro, G. Zagotto, C. Bailly, E. Menta, G. Capranico and M. Palumbo (1998) DNA-binding preferences of bisantrene analogs: relevance to the sequence specificity of drug mediated topoisomerase II poisoning. *Mol. Pharmacol.* **54**, 1036–1045.
22. Maniatis, T., E. F. Fritsch, and J. Sambrook (1982) Molecular cloning: a laboratory manual, pp. 458–469. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Remers, W. A., T. P. Wunz, R. T. Dorr, D. S. Alberts, C. L. Tunget, J. Binspahr and S. Milton (1987) New antitumor agents containing the anthracene nucleus. *J. Med. Chem.* **30**, 1313–1321.
24. Beckett, A. H. and J. Walker (1963) Steric interactions in substituted cyclohexadienes. II. Meso-substituted dihydroanthracenes; steric effects in the reactions of cis and trans isomers. *Tetrahedron* **19**, 545–556.
25. Bissell, R. A., E. Calle, A. Prasanna de Silva, S. A. De Silva, H. Q. N. Gunaratne, J. L. Habib-Jiwan, S. L. A. Peiris, R. A. D. D. Rupasinghe, T. K. S. D. Samarasinghe, K. R. A. S. Sandanayake and J. P. Soumillion (1992) Luminescence and charge transfer. Part 2. (Aminomethyl)anthracene derivatives as fluorescent PET (photoinduced electron transfer) sensors for protons. *J. Chem. Soc. Perkin Trans. 2* **9**, 1559–1564.
26. Naga, K. M., K. J. Snow, B. S. Perrin, J. Thota and C. V. Kumar (2005) Contributions of a long side chain to the binding affinity of an anthracene derivative to DNA. *J. Phys. Chem. B* **109**, 11810–11818.
27. Blandamer, M. J., P. M. Cullis, and B. F. N. Engberts (1998) Titration microcalorimetry. *J. Chem. Soc. Faraday Trans.* **94**, 2261–2267.
28. Rodger, A. (1993) Linear dichroism. *Methods Enzymol.* **226**, 232–258.
29. Naga, K. M., K. J. Snow, B. S. Perrin, M. Duff, A. Bhambhani and C. V. Kumar (2006) *J. Photochem. Photobiol.* **177**, 43–54.
30. Chaires, J. B. (1997) Energetics of drug–DNA interactions. *Biopolymers* **44**, 201–215.
31. Chaudhari, A., J. Thota, and C. V. Kumar (2004) Binding and cleavage studies of two proteins intercalated at the galleries of α-zirconium phosphate. *Micropor. Mesopor. Mater.* **75**, 281–291.
32. Nordén, B., M. Kubista and T. Kurucsev (1992) Linear dichroism spectroscopy of nucleic acids. *Q. Rev. Biophys.* **25**, 51–171.

‡ Preliminary studies showed that when the diethylamine substituents of BDEA are replaced by larger substituents such as triethylamine, or diazabicyclooctane (DABCO), the binding behavior of these ligands with CT DNA was similar to that of BDEA.

33. Nordén, B. and T. Kurucsev (1994) Analyzing DNA complexes by circular and linear dichroism. *J. Mol. Recognit.* **7**, 141–156.
34. Lyng, R., A. Rodger, and B. Nordén (1992) The CD of ligand-DNA systems. 2. Poly(dA-dT) B-DNA. *Biopolymers* **32**, 1201–1214.
35. Lyng, R., A. Rodger, and B. Nordén (1992) The CD of ligand-DNA systems. I. Poly(dG-dC) B-DNA. *Biopolymers* **31**, 1709–1719.
36. Loskotova, H. and V. Brabec (1999) DNA interactions of cisplatin tethered to the DNA minor groove binder distamycin. *Eur. J. Biochem.* **266**, 392–402.
37. LDr in the DNA and probe (anthracycline) region are both negative for intercalation. LDr (260 nm) is about the same magnitude as LDr of probe when intercalated. (Eriksson, M., Norden, B. and S. Eriksson [1988] Anthracycline-DNA interactions studied with linear dichroism and fluorescence spectroscopy. *Biochemistry* 27, 8144–8151.)
38. Classical intercalators, such as ethidium bromide, produce unwinding, elongation, and stiffening of the DNA helix, resulting in increase of LDr magnitude. (Sehlstedt, U., S. K. Kim and B. Nordén [1993] Binding of 4′,6-diamidino-2-phenylindole to [poly(dI-dC)]2 and [poly(dG-dC)]2: the exocyclic amino group of guanine prevents minor groove binding. *J. Am. Chem. Soc.* 115, 12258–12263.)
39. Petraccone, L. (2004) Simultaneous effect of cadaverine and osmolytes on CT-DNA thermal stability. *Thermochim. Acta* **418**, 47–52.
40. Kurnit, B., B. Shafit and J. Maio (1973) Multiple satellite deoxyribonucleic acids in the calf and their relation to the sex chromosomes. *J. Mol. Biol.* **81**, 273–284.
41. Du, W., B. Wang, Z. Li, J. Xiao, G. Yuan, and W. Huang (2003) Interactions of calf thymus DNA with short chain oligoamides. *Thermochim. Acta* **395**, 257–263.
42. Geller, K. and K. E. Reinert (1980) Evidence for an increase of DNA contour length at low ionic strength. *Nucleic Acids Res.* **8**, 2807–2822.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics 326 (2006) 210–220
Chemical Physics
www.elsevier.com/locate/chemphys

Quantitation of protein orientation in flow-oriented unilamellar liposomes by linear dichroism

Jascindra Rajendra [a], Angeliki Damianoglou [a], Matthew Hicks [a], Paula Booth [b], P. Mark Rodger [a], Alison Rodger [a,*]

[a] Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
[b] Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, UK

Received 15 July 2005; accepted 28 February 2006
Available online 18 March 2006

Abstract

The linear dichroism of the visible wavelength transitions of retinal have been used to analyse linear dichroism spectra to determine the orientation of aromatic and peptide structural motifs of Bacteriorhodopsin incorporated into unilamellar soy bean liposomes. The results are consistent with the available X-ray data. This proves that visible light absorbing chromophores can be used to analyse linear dichroism data to give the orientation of membrane proteins in membrane mimicking environments. The work has been extended by screening a wide range of hydrophobic molecules with high extinction coefficients in transitions above 300 nm to find molecules that could be used as independent probes of liposome orientation for experiments involving proteins incorporated into liposomes. Three probes were found to have potential for future work: bis-(1,3-dibutylbarbituric acid)pentamethine oxonol ($DiBAC_4$), retinol and rhodamine B. All three can be used to determine the orientation of the porphyrin of cytochrome *c*, the aromatic residues of gramicidin and the helices of both proteins. The orientation parameter, *S*, for the liposomes varied from batch to batch of unilamellar liposomes prepared by extruding through a 100 nm membrane. The value and variation in *S* was 0.030 ± 0.010. Repeat experiments with the same batch of liposomes showed less variation. Film *LD* data were measured for $DiBAC_4$ and rhodamine B to determine the polarisations of their long wavelength transitions.

Keywords: Liposomes; Linear dichroism; Orientation; Proteins; Probe molecules; Bacteriorhodopsin

1. Introduction

Processes taking place on or in lipid bilayer membranes are essential for biological systems, for example, for translocation of molecules into and out of cells and membrane-bound receptor activation. The understanding of inter-molecular interactions within the membrane, between the lipids themselves, between lipids and other species (proteins, drugs, and ions), and between membrane-bound molecules are all important. With the majority of future drug targets being anticipated to be membrane proteins, knowing the locations and orientations of molecules bound to the membrane is an extremely important goal. However, to date no simple experimental method has been established to achieve this for membrane-bound proteins and little work has been done with small molecules. Ardhammar et al. [1] showed that liposomes (spherical bilayer systems) could be oriented by Couette flow and then linear dichroism (*LD*, the differential absorbance of light polarised parallel and perpendicular to an orientation direction) used to probe the orientation of small molecules in membranes. We extended this work by showing that the orientations of proteins both in (e.g. gramicidin) and on the surface of membranes (e.g. a thylakoid membrane protein, pre-PsbW, and cytochrome *c*) could be probed qualitatively using Couette flow linear dichroism [2,3]. Although *LD* only gives an

* Corresponding author. Tel.: +44 24 76574696; fax: +44 24 76524112/75795.
E-mail address: a.rodger@warwick.ac.uk (A. Rodger).

0301-0104/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2006.02.036

average orientation, whereas other techniques such as X-ray or NMR can give specific information about the atoms in the structure, the *LD* method has the great advantage of being able to be readily performed on almost any system one might wish to study.

The reduced linear dichroism, LD^r, of a chromophore oriented in a liposome has been shown to be [1,3]:

$$LD^{\mathrm{r}} = \frac{LD}{A} = \frac{A_{\parallel} - A_{\perp}}{A} = \frac{3S}{4}(1 - 3\cos^2\beta) \quad (1)$$

where *LD* is the measured *LD* signal, *A* is the isotropic absorbance of the same sample in a cell of the same pathlength, $A_{\parallel}$ is the absorbance of light polarised parallel to the orientation direction, *S* is the so-called orientation parameter (= 1 for full orientation and 0 for random orientation) and β is the angle between the chromophore's transition moment and the normal to the liposome surface (i.e. the average lipid direction) (Fig. 1). The quantitative interpretation of the *LD* spectra of membrane-bound proteins following, e.g., Ref. [4] has proved challenging due mainly to two factors: (i) the sloping baseline introduced by light scattering and (ii) the fact that we have not been able to tell the degree of orientation of the liposomes [1–3,5]. Ardhammar et al. used sucrose to match the refractive indices of the liposomes and the aqueous solution around the liposomes to reduce the light scattering. In our newly developed capillary Couette flow *LD* cell [6], whose light scattering is intrinsically lower than that of traditional Couette cells, we found this did not improve the signal to noise and at wavelengths below 210 nm the sucrose absorbance precluded measurements – thus rendering this approach inapplicable to proteins.

Our capillary *LD* cell is based on a rotating quartz capillary, sealed at one end, as the sample holder. A stationary quartz rod is inserted into the capillary to create the shear force and focusing lenses before and after the sample ensure the light is incident on the sample and then photomultiplier tube [6]. With this cell we have significantly reduced the scattering artefacts in a range of samples including fibrous proteins, liposomes and carbon nanotubes [7–10]. We have also found that further progress can sometimes be made following the methods of Nordén, Nordh, and Mikati [11] to correct for light scattering. Thus, while the issue of light scattering with liposome samples is still a problem, it is no longer dominant.

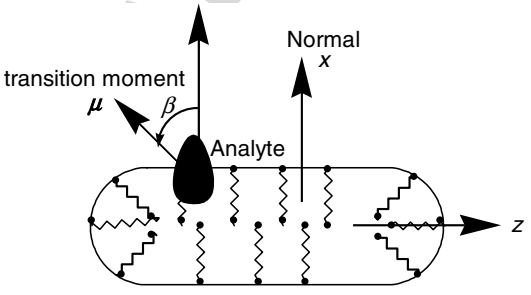


Fig. 1. Schematic illustration of an analyte in a shear-deformed liposome.

A variety of probe molecules and different techniques have been used to determine the lipid vesicle orientation factor. Ardhammar et al. used light scattering of flow-oriented liposomes to gauge the degree of orientation [1]. They also used $Ru(phen)_2dppzcpCOOCH_3$ as a probe molecule to determine the *S* factor [12]. A further estimate of degree of orientation of the membrane was obtained by Brattwall et al. using penetratin [13]. Castanho and Lopes used Nystatin A_1 and Amphotericin B which belong to a polyene antibiotic family as probe molecules to study the orientational distribution function in a lipid bilayer film [14]. None of these approaches have been developed into a routine method for assessing the orientation of a protein molecule in cell-membrane mimicking environments.

The focus of the work reported in this paper is therefore to address the question of 'what is the orientation parameter of the liposomes?' The ideal would be an internal 'standard that can be added to all samples to give a direct read-out of the orientation parameter, in much the same way as DNA bases can be used for determining the orientation on DNA of ligands, e.g. [4].

In our work, we have considered a range of liposome preparation methods and a series of probe molecules, aiming to find one to quantify the liposome orientation and consequently the protein orientation in the liposomes: if a probe molecule of known transition polarisation is inserted into liposomes, using Eq. (1) its *LD* could be used to estimate the *S* factor of the shear-deformed liposomes, and consequently the angle β between the protein's (or any other analyte's transition moments) and the lipids could be determined. Such a probe molecule must have a number of characteristics: it should bind to the membrane (rather than stay free in the aqueous solution); it should absorb outside the protein's absorption region (and very little within it); it should have a high extinction coefficient; it should not interact with the protein; and it should have the least possible effect on the protein's *LD* signal.

2. Materials and methods

2.1. Materials

All reagents, including soy bean lipid (L-α-Lecithin, Type IV-s: from soybean approx. 40% (TLC)), Gramicidin D (from *Bacillus brevis*) and bacteriorhodopsin (from *Halobacterium salinarium* strain S9, lyophilised powder) were obtained from Sigma Chemical Company and were used as received except for bis-(1,3-dibutylbarbituric acid)pentamethine oxonol ($DiBAC_4$) which was purchased from Molecular Probes; undecylprodiginine which was kindly provided by Dr. G.L. Challis and Dr. Christophe Corre, University of Warwick; methanol and ethanol (analytical grade) which were obtained from BDH laboratory supplies; and chloroform which was purchased from Fisher Scientific. Polycarbonate membranes of 100 nm pore size were purchased from Avestin Inc.

2.2. Equipment

UV–Visible absorbance spectra were recorded using a Jasco V-550 or Cary IE spectrophotometer and circular dichroism (*CD*) with a Jasco J-715 spectropolarimeter using a 1 mm pathlength quartz cell. Flow linear dichroism spectra were recorded using a Jasco J-715 circular dichroism spectropolarimeter, which was adapted for flow *LD* measurements, and a 0.5 mm pathlength quartz capillary *LD* cell [6]. (Note the factor of two the difference in absorbance and *LD* pathlengths requires in any calculation of LD^r.) The *LD* cell is a cylindrical capillary flow cell, consisting of two coaxial cylinders, one of which is a capillary and the other is a quartz rod. There is a 0.25 mm annular gap between the rotating capillary and the stationary rod where the solution of interest is subjected to a constant flow, and hence viscous drag, by rotating the capillary at a speed large enough to cause significant orientation but not a turbulent flow (in our case 3000 rpm).

Absorbance and *CD* baselines were collected in the same cuvette using a liposome preparation identical to that of the corresponding sample but without the analytes. For *LD*, the best baseline to subtract from a recorded spectrum (as judged by the least scattering baseline in the final spectrum) was usually that of the sample in the same cell but without rotation. Alternatively, a spectrum of a rotating liposome solution was used as the baseline.

2.3. Liposome preparation

In this work liposomes were prepared by three different methods: the vortex method, the liquid-nitrogen method and the sonication method.

1. *Vortex method*: Aqueous buffer solution (e.g. 4 mL of 5 mM phosphate buffer, pH 7) was added to L-α-phospholipidycholine lipids (e.g. 10 mg) to obtain a lipid concentration of 2.5 mg/mL. The sample was votexed to obtain multilamellar liposomes and then extruded 11 times using a Liposofast Basic Extruder (Glen Creston Ltd.) with a polycarbonate membrane (100 nm pore size) to obtain unilamellar liposomes [15].
2. *Nitrogen method*: Chloroform (2 mL) or chloroform and methanol (2 mL and 1 mL, respectively) were added to L-α-phospholipidycholine lipids (2.5 mg) and left overnight under a stream of nitrogen. The resulting dry lipid film was vortexed with phosphate buffer solution (1 mL, 5 mM, pH 7) to obtain multilamellar liposomes, then extruded 11 times using a Liposofast Basic Extruder with a polycarbonate membrane (100 nm pore size) to obtain unilamellar liposomes, and finally frozen using liquid nitrogen (−196 °C) and thawed under running tap water (4 times).
3. *Sonication method*: L-α-Phospholipidycholine lipids (2.5 mg) were placed in a vial with a mixture of organic solvents and was sonicated for 10 s at ambient temperature. The sample was left overnight under a stream of nitrogen. The resulting dry lipid film was vortexed with 1 mL of phosphate buffer (5 mM, pH 7).

2.4. Analyte introduction

The probe molecules and proteins were sometimes introduced to liposome solutions by adding dropwise an aqueous buffered liposome solution (4 °C) to a concentrated solution of the analyte (also at 4 °C) dissolved in an appropriate (water miscible) solvent. For samples containing probes and proteins, they were usually added in this order. Cytochrome *c* was added as a solid. Alternatively, the nitrogen method was used, with the analyte being added to the organic solution of lipid. If required, samples were subsequently diluted by adding either buffer or aqueous solutions of unilamellar liposomes. In most cases, samples were left overnight to equilibrate.

2.5. Film LD sample preparation

Film *LD* samples were prepared [4] by dissolving (heating to near boiling) low molecular weight polyvinyl alcohol (PVA) (0.48 g) in water (4.8 mL). For the sample films the viscous solution was allowed to cool before a saturated methanolic solution of the probe (0.2 mL) at an appropriate concentration (e.g. 0.1 mM for rhodamine B) was added. The mixture was cast onto a glass plate and allowed to dry in the dark over a period of two days. The dry film was removed from the plates with the aid of a scalpel and stretched in a mechanical film stretcher by a factor of two under heat from a hair dryer (~60 °C). *LD* and normal absorption were recorded where the wavelength range and data interval of the two instruments were set to correspond. A control film was prepared by adding 0.2 mL of methanol in the place of probe solution.

3. Results and discussion

LD data for bacteriophodopsin which has retinal as a convenient internal probe are described first followed by the results for a range of potential small molecule probes for determining the orientation of liposomes in a Couette flow cell. Finally, some data from selected probes together with the model proteins gramicidin and cytochrome *c* are described.

3.1. Bacteriorhodopsin orientation in liposomes

Bacteriorhodopsin (BR) is found in the purple membrane of Halobacteria that live in salt marshes [16,17], it is a 248 residue protein of mass 26,000 Da and includes a retinal chromophore (Fig. 2) covalently bound to a lysine. Each BR has 7 transmembrane helices, 3 of which in the crystal have their axis at ~70° to the lipids and the remaining 4 are parallel to the lipids [18–21]. The long axis of the

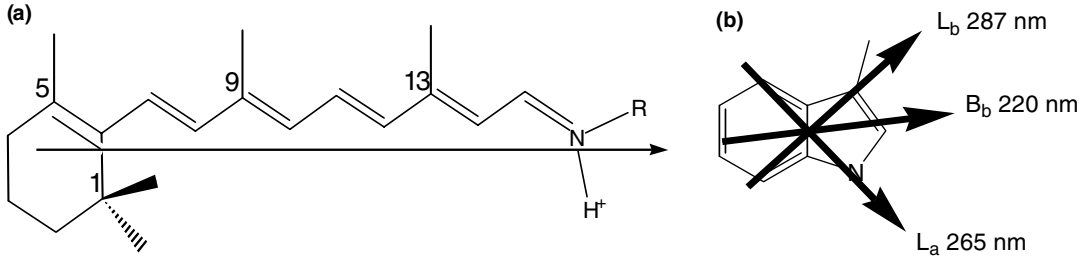


Fig. 2. (a) All-*trans* retinal converted to the Shiff base. The vector represents the transition dipole moment of the 570 nm transition of retinal [18]. (b) Tryptophan and its transition polarisations [23].

retinal lies at ~69° to the lipids in the crystal. An FTIR study on a dried suspension of BR in lipids gave an average angle of 90 ± 20° between the retinal plane and the plane of the purple membrane with the C_9=C_{10} (Fig. 2) double bond of retinal being nearly parallel to the plane of the membrane (this makes it ~30° between the retinal long axis and the average lipid). BR is known to insert into liposomes in a single orientation [22].

The absorption, *CD* and *LD* spectra of bacteriorhodopsin in liposomes are given in Fig. 3. The absorption spectra (Fig. 3(a)) are consistent with those reported in the literature for the initial state of BR [24]. The absorption maximum at ~570 nm is due to a long axis polarised transition of the retinal chromophore. The broad peak in the near UV region (260–290 nm) is due to the transitions of the protein aromatic side chains phenylalanine, tyrosine and tryptophan and the peak observed in the far UV region (215 to ~230 nm) is due to the peptide $n \rightarrow \pi^*$ transition of the amide groups. The backbone *CD* spectrum of BR (Fig. 3(b)) is in accord with the *CD* spectrum for a highly α-helical protein as expected [25,26].

A peak at ~570 nm is observed in the *LD* spectrum of BR (Fig. 3). Taking the literature value of the orientation of the long axis of retinal with respect to the lipids of a lipid bilayer (69°)[96] and the value for $LD^r_{570\ nm}$ of 0.022 (from the Fig. 3(a) spectra, noting that the absorbance pathlength is twice that of the *LD* cell), it follows from Eq. (1) that $S \sim 0.048$. This value varies by about 5–10% for independently prepared samples with the same batch of liposomes and by up to 30% for different batches of the same average size liposomes.

The aromatic region (260–280 nm) of the *LD* spectrum is dominated by the indole chromophore of the tryptophan (W) residues [23]. Contributions from L_a (270 nm) and L_b (287 nm) (Fig. 3(a)) transitions are apparent, both showing positive *LD*. Using $S \sim 0.048$ and $LD^r_{287\ nm} = 0.0076$ it follows from Eq. (1) that $\beta(L_a) \sim 60°$. Similarly, $\beta(L_b) \sim 65°$. Thus, the *LD* indicates that the average W is tilted so that the normal to the plane of the W is ~40° from the average lipid and the long and short axes are both at an angle of 60–65° to the lipids. This is consistent with the fact that the retinal is sandwiched by tryptophan residues in the X-ray structure [25].

The protein backbone *LD* spectrum shows a positive maximum at 220 nm ($n \rightarrow \pi^*$) and a negative maximum at ~213 nm ($\pi \rightarrow \pi^*$). Again using $S \sim 0.048$, it follows that the $n \rightarrow \pi^*$ transition (which is polarised perpendicular to the α-helix long axis) is at ~58° [3] ($LD^r_{220\ nm} = 0.006$) from

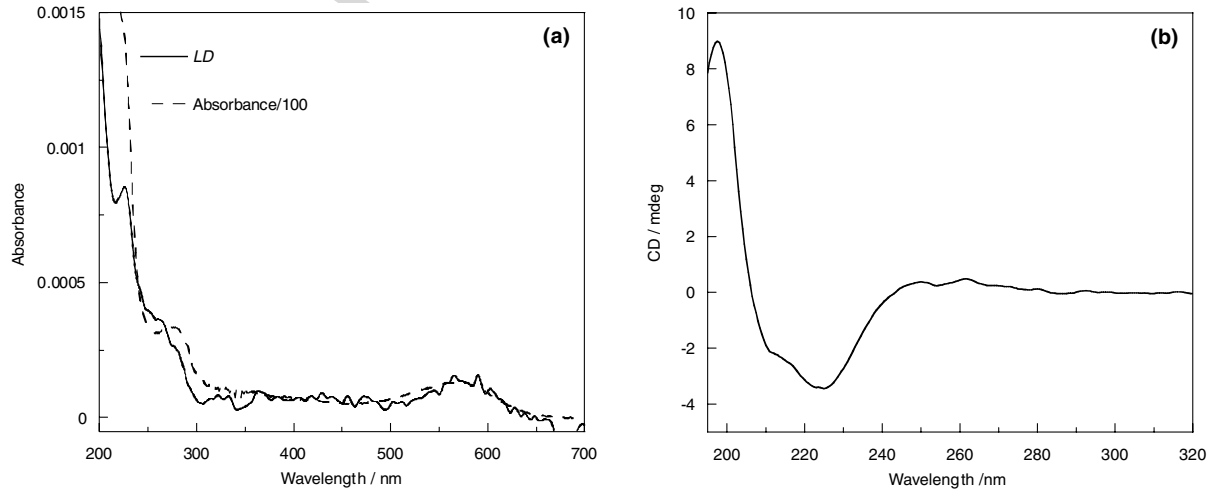


Fig. 3. Spectra of bacteriorhodopsin (0.2 mg/mL) added to a liposome solution (0.5 mg/mL): (a) Absorption (dashed line, 1 mm pathlength, baseline: liposome absorption spectrum) and *LD* (solid line, 0.5 mm pathlength, baseline: *LD* spectrum of sample without rotation spectrum); (b) *CD* (1 mm pathlength, baseline: liposome CD spectrum).

the average lipid direction. Thus, the average orientation of the transmembrane helices is ~30° from the membrane normal (Fig. 4). Our average value suggests that the protein is less rigidly held in a liposome than when dried or crystallised, which is entirely reasonable. Thus, we can conclude that a covalently bound internal probe of a membrane protein can be used semi-quantitatively to analyse the orientation of protein chromophores in liposomes.

3.2. Independent probes to determine liposome orientation

Very few proteins have unique spectroscopically isolated internal probes of known orientation with respect to the rest of the protein, such as retinal in BR. Having proved the concept of using a chromophore with visible spectroscopy to probe protein motif orientation our aim was therefore to find one or more probe molecules that could be added to a liposome preparation to give an orientation read-out for other proteins. The holy grail is a molecule with no spectroscopic signature in the protein region, 100% membrane affinity (so there is no probe in the aqueous part of the solution contributing to the absorbance but not the *LD*), no precipitation from an aqueous liposome solution, and no perturbing effect on the membrane environment.

Potential probe molecules were screened for their ability to give an *LD* spectrum when added to an aqueous liposome solution following any of the preparation methods outlined above. From over 30 candidates, 7 (Fig. 5) were selected for further study on the basis of our ability to get them into solution and bound to liposomes: Amphotericin B (AmB), hemin, undecylprodiginine, $DiBAC_4$, retinol, protoporphyrin and rhodamine B. AmB, a polyene antibiotic used in the medical treatment of fungal infections [27], was prepared by adding an aqueous liposome solution dropwise to a concentrated DMSO solution of AmB. Hemin, whose visible absorbance has two non-degenerate components of the Soret band at ~400 nm, was similarly prepared from a concentrated methanolic solution as was the red coloured antibiotic undecylprodiginine (whose linear form was obtained from a cell extract using HPLC [28]). Bis-(1,3-dibutylbarbituric acid)pentamethine oxonol ($DiBAC_4$), which is used as a probe for ion channels and carriers in lipid membranes, has a very high extinction coefficient (160,000 mol dm^{-3} cm^{-1}) in the visible [29]. It was simply added to liposome solutions from methanol. Retinol was dissolved in isopropanol. Protoporphyrin samples were prepared by the nitrogen/freeze–thaw method as it was not sufficiently soluble in a water miscible solvent. rhodamine B samples were also prepared in this manner as this method seemed to give better spectra.

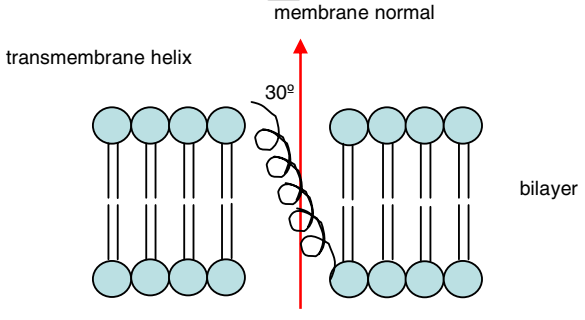


Fig. 4. Schematic illustration of the average orientation of the transmembrane helices of BR in the liposome bilayer deduced from the *LD*.

The absorbance and *LD* of each of these probes when in solution with unilamellar soy bean liposomes are given in Fig. 6. AmB, hemin, undecylprodiginine and protoporphyrin were eliminated from our search for the ideal probe for different reasons. AmB absorbance and *LD* signals in liposome solution are not proportional to its concentration in the solution. Gruszecki et al. have previously concluded that AmB adopts two different orientations when bound to a lipid bilayer [27] and we presumably are seeing a concentration dependence of the occupancy of the two binding modes. In any case it renders AmB unsuitable for use as a generic probe. However, AmB does have a sharp peak at 325 nm and a fairly flat spectrum between 200 and 300 nm. In addition it does co-incorporate into the membranes with proteins (data not shown). So there may be instances where it could be used. Hemin proved to absorb too much at 200 nm to be suitable and protoporphyrin has a complicated spectrum with significant intensity below 300 nm. Undecylprodiginine was also eliminated because it stuck to the quartz of cuvettes and *LD* cell capillaries. The remaining 3 probes: $DiBAC_4$, retinol and rhodamine B have *LD* intensities proportional to their concentration (Fig. 7). Their LD^r spectra are given in Fig. 6.

$DiBAC_4$, retinol and rhodamine B all have potential as liposome orientation probes. To determine how they are oriented on or in the liposomes one needs to know their own transition polarisations. Retinol is available from the literature: its long wavelength visible transition is along its long axis as illustrated in Fig. 2. For $DiBAC_4$ and rhodamine B it was necessary to measure film *LD* spectra (Fig. 8). For $DiBAC_4$ in a PVA film, $LD^r = 0.48$ at 509 nm and 0.83 at 602 nm (Fig. 8). In a film, $DiBAC_4$ can be treated as a uniaxially oriented 'rodlike' molecule [4]. So

$$LD^r = \frac{LD}{A} = \frac{3}{2}S(3\cos^2\alpha - 1) \tag{2}$$

where S is the orientation parameter and α the angle between the orientation direction and the transition moment of interest. We assume the maximum LD^r is related to the orientation parameter according to

$$LD^r_{\max} = 3S_z \tag{3}$$

As $DiBAC_4$ transition polarisations are almost determined by symmetry to be either long or short axis polarised, it is reasonable to conclude that the 602 nm transition lies along the long axis of $DiBAC_4$ as it has the largest LD^r signal. This gives $S = 0.28$, or more strictly, $S \geqslant 0.28$ (allowing for the fact the molecule is not quite of $\mathbf{D}_{2h}$ symmetry). It

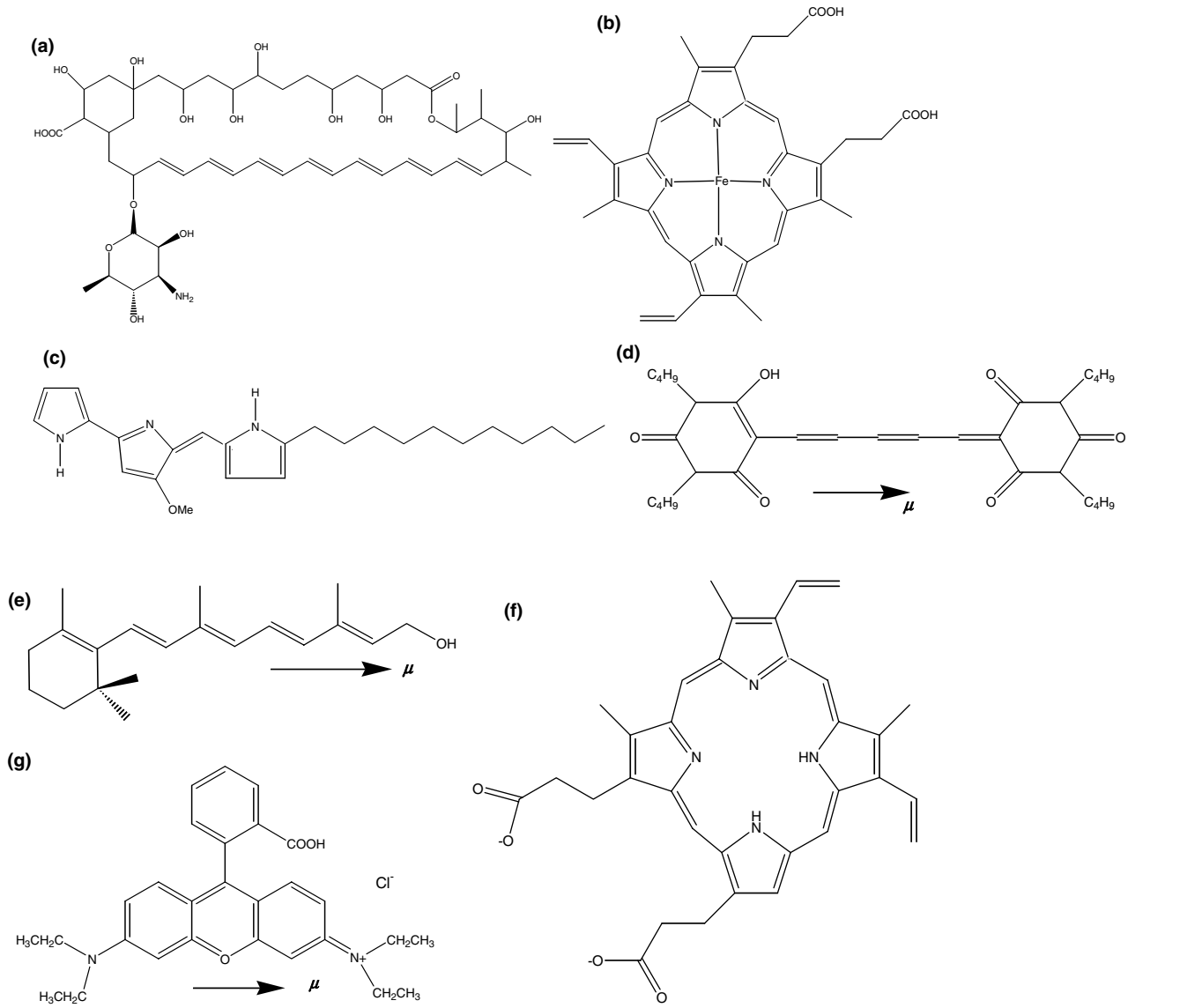


Fig. 5. (a) Amphotericin B (AmB), (b) hemin, (c) undecylprodiginine, (d) $DiBAC_4$, (e) retinol, (f) protoporphyrin and (g) rhodamine B. Arrows indicate the polarisations of the long wavelength transitions.

then follows that the 509 nm region of the spectrum is polarised $\geqslant 32°$ from the long axis of $DiBAC_4$. Given the approximate symmetry of the molecule and also the region of the spectrum, this means that this is the result of overlapping transitions of different polarisations. The rhodamine B chromophore similarly has approximate $\mathbf{C}_{2v}$ symmetry so it is reasonable to conclude that the 580 nm component of its 550 nm band (largest LD^r) is along its long axis and the 540 nm component is perpendicular to this. (Although the film *LD* signal is not negative at 540 nm, if the actual component spectra are determined using Thulstrup's trial and error method it is apparent that rhodamine B's absorbance is short axis polarised at 540 nm [4]).

Using the transition polarisation information and the liposome *LD* spectra of Fig. 6, we can then determine the orientation of the probes in the liposomes and hence the liposome orientation parameters. $DiBAC_4$ loses the mixed polarisation 509 nm band in the lipid environment, but the 602 nm band seems unaffected. The strong positive *LD* of this transition in the liposomes says it is well-oriented perpendicular to the average lipid, perhaps lying on the liposome surface with its alkyl chain inserting in. $LD^r_{600\ \mathrm{nm}} \sim 0.019$, which from Eq. (1) gives: $S \sim 0.026$, assuming $\beta_i = 90°$. The rhodamine B in the liposome *LD* spectrum of Fig. 6 has a similar profile to that of the film data (Fig. 8) suggesting that the long axis of the probe is lying parallel to the surface and the short axis parallel to the lipids (cf. Eq. (1)). If this is the case, it then follows from the 580 nm LD^r value (which is not overlaid with any other transition) of 0.025 that *S* for the liposomes $\sim$0.032. Retinol, assuming $\beta = 90°$, yields $S = 0.020$. Allowing for any flexibility within the membrane and a less rigid orientation results in a decrease in β and an increase value for *S*. Repeat experiments with all three probes, lead

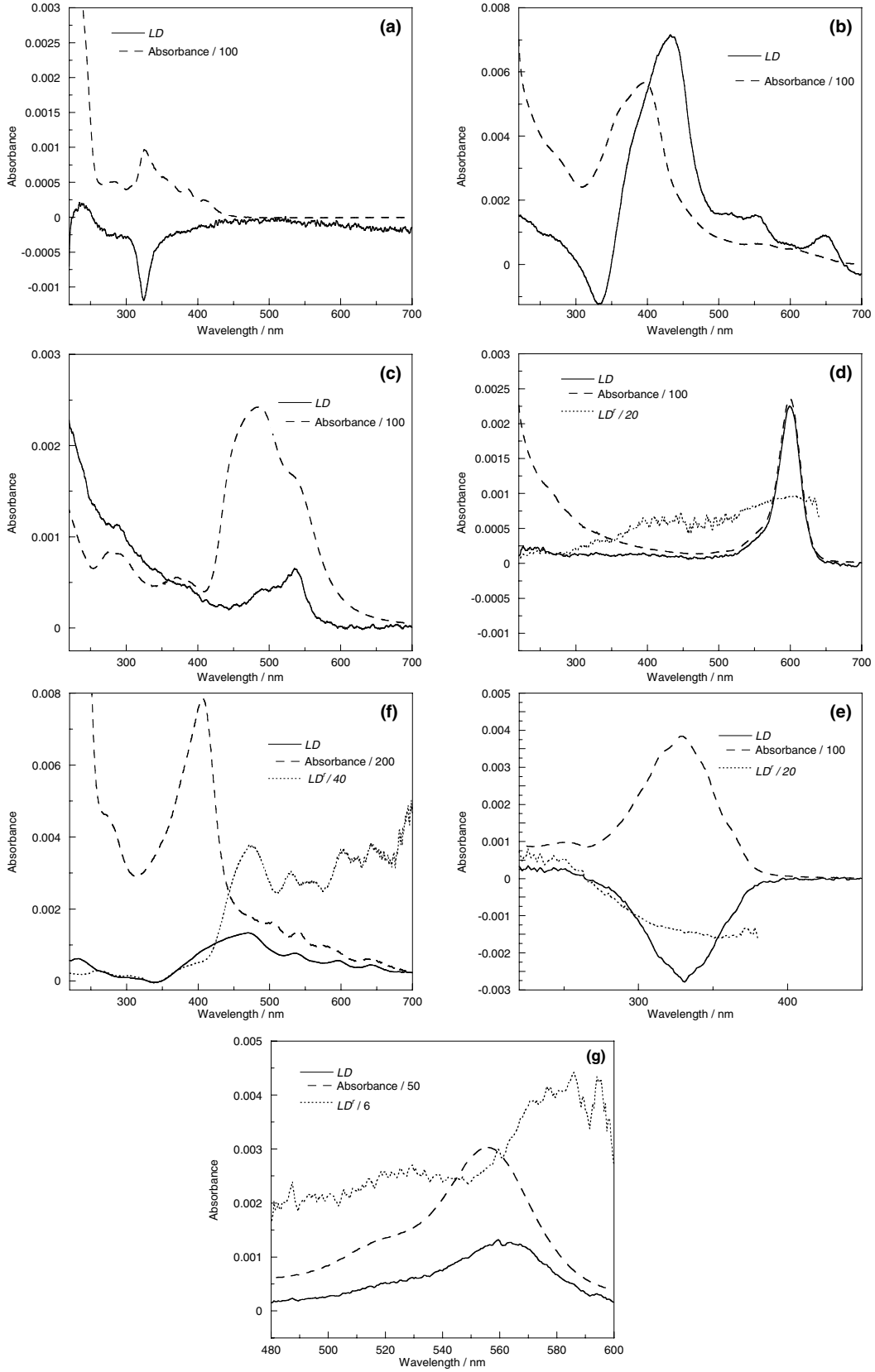


Fig. 6. Absorbance, *LD* and in some cases LD^r (note *LD* cell pathlength is half that of the normal absorbance cell used) of: (a) AmB (0.02 mg/mL) in aqueous liposome solution (5 mg/mL); (b) hemin (0.5 mg/mL) in aqueous liposome solution (2.5 mg/mL); (c) undecylprodiginine (0.5 mg/mL) in aqueous liposome solution (1 mg/mL); (d) $DiBAC_4$ (0.01 mg/mL) in aqueous liposome solution (4 mg/mL); (e) retinol (0.05 mg/mL) in aqueous liposome solution (4.5 mg/mL); (f) protoporphyrin (0.1 mg/mL) in aqueous liposome solution (0.5 mg/mL); (a) rhodamine B (0.01 mg/mL) in aqueous liposome solution (0.5 mg/mL). A soybean lipid spectrum was used as blank for absorbance and sample without rotation used as blank for *LD* in each case.

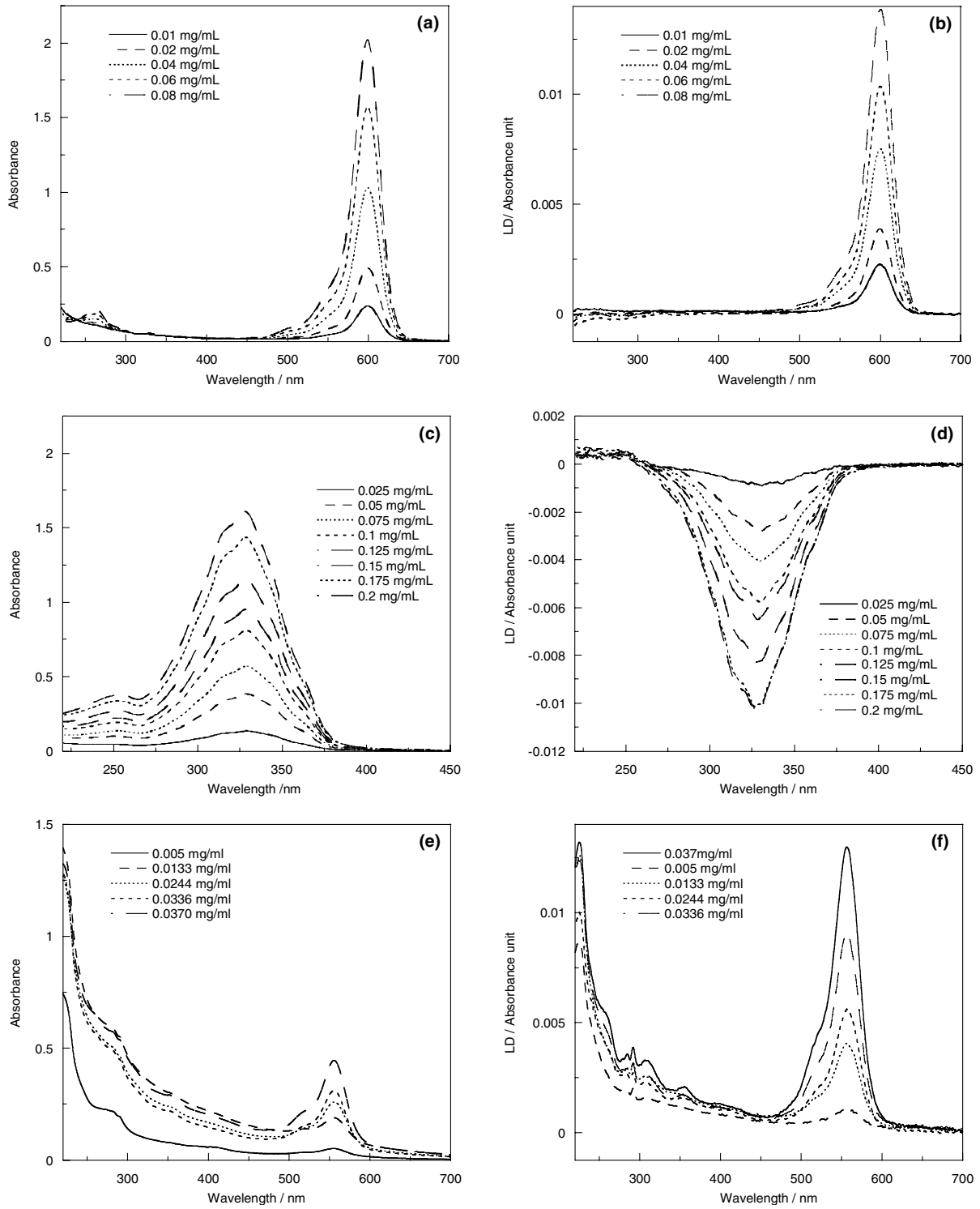


Fig. 7. Absorbance and *LD* spectra of different concentrations of selected probes (a), (b) $DiBAC_4$ (liposome concentration: 4 mg/mL); (c), (d) retinol (liposome concentration: 4.5 mg/mL); (e), (f) rhodamine B (liposome concentration: 2.5 mg/mL) in aqueous liposome solutions. Probe concentrations are indicated in the figures. Absorbance and *LD* spectra are base line corrected by subtracting liposome (4 mg/mL) absorbance and *LD*, respectively.

us to conclude that the liposomes are oriented in our experiments with an orientation parameter of $S \sim 0.030 \pm 0.010$. Independently prepared sample using the same batch of liposomes had a 5–10% variation in value of *S*. Different batches of liposomes resulted in the wider range.

3.3. Probe molecules and proteins

We have previously shown that both gramicidin and cytochrome *c* can be oriented in Couette flow and give *LD* spectra when in the presence of liposomes. So we were

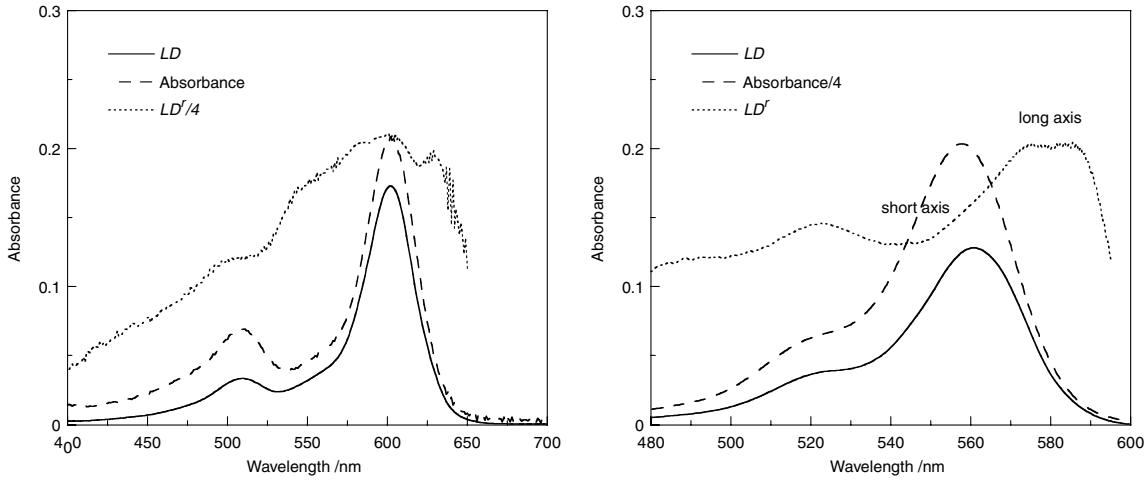


Fig. 8. *LD*, absorbance and LD^r of: (a) $DiBAC_4$ and (b) rhodamine B in stretched PVA films.

interested to see whether the probe molecules could be used to make the geometric information deduced more quantitative. Gramicidin is a linear polypeptide with an alternating sequence of L- and D-aminoacids [30]: formyl-L-Val–Gly–L-Ala–D-Leu–L-Ala–D-Val–L-Val–D-Val–L-Trp–D-Leu–L-Trp–D-Leu–L-Trp–D-Leu–L-Trp-ethanolamine. We have previously shown that using our preparation methods gramicidin forms head to tail dimers in liposomes [31,3]. Cytochrome *c*, by way of contrast, oriented in such a way as to be consistent with it being bound on the surface of the membranes [3]. So the two proteins are simple systems to let us address the question of whether the probe affects protein membrane binding.

3.4. $DiBAC_4$

Fig. 9(a)–(d) shows that both proteins and $DiBAC_4$ can be simultaneously oriented. The degree of protein orientation does not seem to be affected by $DiBAC_4$, however, intriguingly, the light scattering is reduced in its presence. This is not due to the reduction in the size of the liposomes which have an average length of 160 nm in each case (data not shown). It may represent a loss of a small population of very large liposomes, or perhaps it is due to the probes facilitating the solubilisation of the protein – particularly in the case of gramicidin. The data of Fig. 9(b) show the *LD* of the porphyrin group of cytochrome *c*. Eq. (1) then leads to a value of ~50° for the angle between the average lipid and the porphyrin plane which is consistent with our expectations based on a perpendicular attachment to the liposome surface.

In dealing with the protein backbone region it is important to ensure that the photomultiplier tube is responding correctly to the signal by diluting the sample by e.g. a factor of 2 and ensuring the signal halves in magnitude and the wavelengths of the maxima stay the same. This was done for the systems reported in this paper (data not shown) and it was confirmed that the samples obeyed the Beer–Lambert Law. Geometry calculations can then be undertaken. For example, the *S* factor for gramicidin (~0.2 mg/mL) incorporated $DiBAC_4$ (0.01 mg/mL)/liposomes (~1.5 mg/mL) system is 0.021. It follows that the angle between the $n \rightarrow \pi^*$ transition (which is perpendicular to the axis of the helix) at 228 nm and the lipid normal is ~70° suggesting that the helices of gramicidin are on average tilted from the lipid normal.

3.5. Retinol

Retinol does not affect the shape of the *CD* spectra of either gramicidin or cytochrome *c*, but its absorbance is broad and sufficiently close to where the protein absorbs that the effect it has on reducing the sample scattering (as in the case of $DiBAC_4$) affects the protein's apparent *LD* spectra (e.g. Fig. 9(f)). Despite this, if one uses retinol as an internal probe to determine orientations for the cytochrome *c*, porphyrin or the gramicidin tryptophans, the results are the same as those determined with $DiBAC_4$ – which is encouraging.

3.6. Rhodamine B

The high extinction coefficient of rhodamine B makes it an attractive probe as it can be used at low concentrations. Its disadvantage is that the best data are collected using the nitrogen preparation method which is more laborious than simply adding lipids to a solution of the probe. Analysis of the orientations of the protein chromophores in the visible, near UV and far UV regions (data not shown) all indicate that it can be used as a probe.

4. Conclusions

The goal of this research was to show that *LD* could be used more quantitatively to determine the orientation of proteins in membrane mimicking systems. The proof of

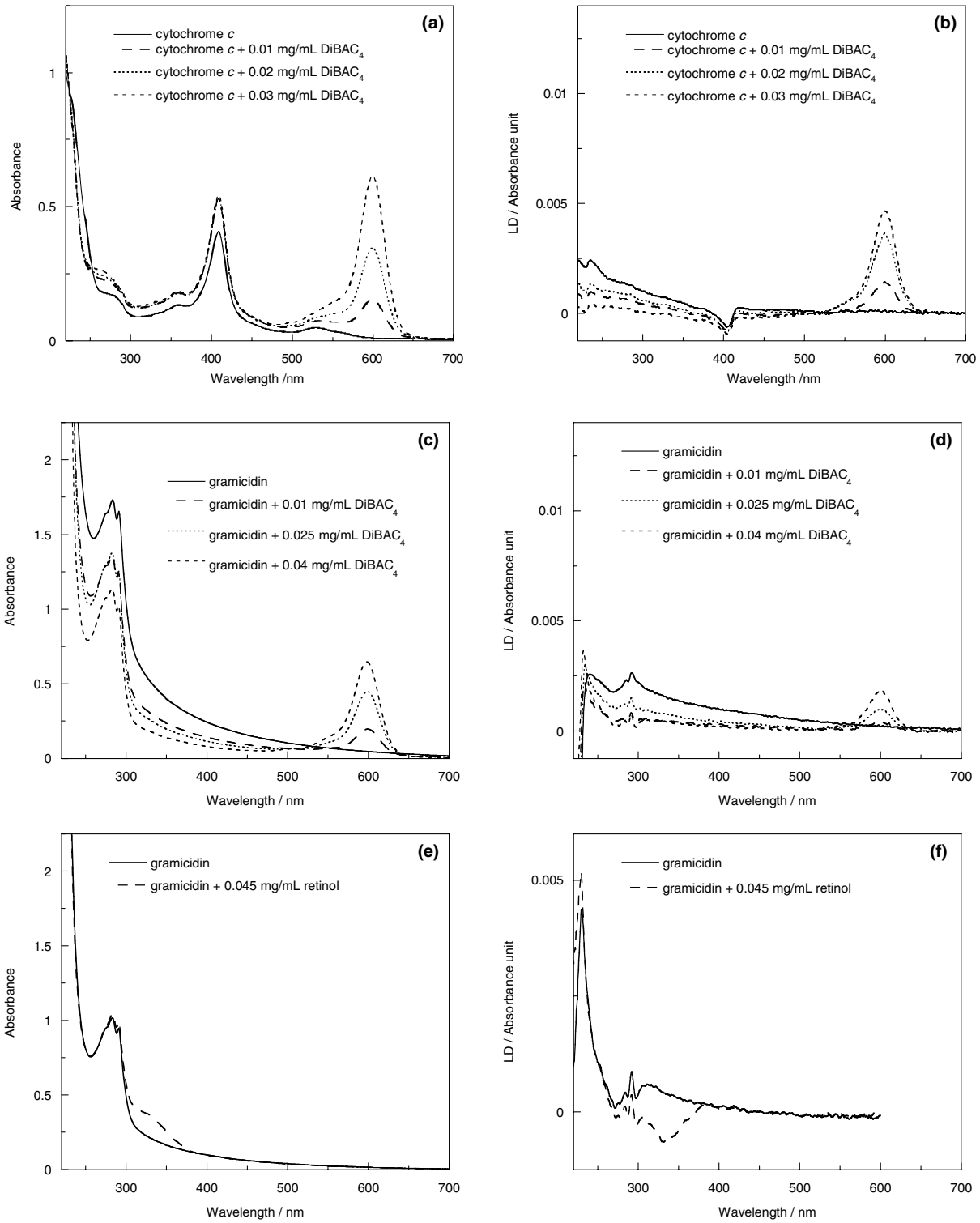


Fig. 9. (a) Absorbance and (b) *LD* of cytochrome *c* (0.6 mg/mL) in a $DiBAC_4$/lipid (5 mg/mL) system. (c) Absorbance and (d) *LD* of gramicidin (0.8 mg/mL) in a $DiBAC_4$/lipid (5 mg/mL) system. $DiBAC_4$ concentrations are indicated on the figures. (e) Absorbance and (f) *LD* of gramicidin (0.6 mg/mL) in a retinol/lipid (5 mg/mL) system. Retinol concentrations are indicated on the figures. Baselines were corrected by subtracting spectra of liposome solutions.

concept experiment was achieved with bacteriorhodopsin which has its own internal probe, retinal, covalently bound to it and for which X-ray data on relative orientations of chromophores are available. For the more generic problem of a protein without a covalently bound probe, a wide range of hydrophobic molecules were screened to find some molecules that had high membrane affinity, little spectroscopy below 300 nm but also have a large absorbance above 300 nm which could be used to determine the orientation of the liposomes. Three probes were found to have potential

for future work: bis-(1,3-dibutylbarbituric acid)pentamethine oxonol ($DiBAC_4$), retinol and rhodamine B. Film *LD* data were measured for $DiBAC_4$ and rhodamine B to determine the polarisations of their long wavelength transitions. The orientation parameter, *S*, varied from batch to batch of liposomes but for unilamellar liposomes prepared by extruding through a 100 nm membrane, $S \sim 0.030 \pm 0.010$. Bacteriorhodopsin with its high protein load had a slightly higher value of *S*. The values show quite a high degree of orientation in flow of the originally spherical liposomes. The orientation parameter for a typical very long DNA sample is ~0.03–0.1 [32]. A single preparation of liposomes showed much less variation. When both probe molecules and proteins were added to liposomes it was generally possible to use the probes to determine the liposome orientation and hence to determine the protein orientation within the membranes. At the low wavelengths relevant for protein backbone motifs, it is important to ensure that the photomultiplier tube is responding properly by measuring on a series of samples of different concentration. If the wavelengths of *LD* maxima remain the same and intensities follow the Beer–Lambert Law then data can be used. The probes seemed to have the general effect of reducing light scattering.

So in conclusion a significant step forward has been taken towards being able to use flow linear dichroism as a quantitative structural tool for proteins in fully hydrated membrane environments such as liposomes.

Acknowledgements

The gift of undecylprodiginine from Dr. G.L. Challis and Dr. Christophe Corre, University of Warwick, was a valuable contribution to this work and EPSRC (GR/T09224/01) and BBSRC (REI20503) are gratefully acknowledged for providing funding. The careful work of a referee is also gratefully acknowledged.

References

[1] M. Ardhammar, N. Mikati, B. Nordén, Journal of the American Chemical Society 120 (1998) 9957.
[2] A. Rodger, J. Rajendra, R. Marrington, R. Mortimer, T. Andrews, J.B. Hirst, A.T.B. Gilbert, R. Marrington, D. Halsall, T. Dafforn, M. Ardhammar, B. Nordén, C.A. Woolhead, C. Robinson, T. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez, M.J. Hannon, Flow oriented linear dichroism to probe protein orientation in membrane environments, in: R.H. Templer, R. Leatherbarrow (Eds.), Biophysical Chemistry: Membranes and Proteins, The Royal Society of Chemistry, Cambridge, 2002, pp. 3–19.
[3] A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J.D. Hirst, A.T.B. Gilbert, T.R. Dafforn, D.J. Halsall, C.A. Woolhead, C. Robinson, T.J. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez, M.J. Hannon, Physical Chemistry Chemical Physics 4 (2002) 4051.
[4] A. Rodger, B. Nordén, Circular Dichroism and Linear Dichroism, Oxford University Press, Oxford, 1997.
[5] M. Ardhammar, P. Lincoln, B. Nordén, Proceedings of The National Academy of Science of USA 99 (2002) 15313.
[6] R. Marrington, T.R. Dafforn, D.J. Halsall, A. Rodger, Biophysical Journal 87 (2004) 2002.
[7] T.R. Dafforn, A. Rodger, Current Opinion in Structural Biology 14 (2004) 541.
[8] R. Marrington, E. Small, A. Rodger, T.R. Dafforn, S. Addinall, Journal of Biological Chemistry 47 (2004) 48821.
[9] J. Rajendra, M. Baxendale, L.G. Dit Rap, A. Rodger, Journal of the American Chemical Society 126 (2004) 11182.
[10] J. Rajendra, A. Rodger, Chemistry–A European Journal (2005).
[11] N. Mikati, J. Nordh, B. Nordén, Journal of Physical Chemistry 91 (1987) 6048.
[12] M. Ardhammar, P. Lincoln, B. Nordén, Journal of Physical Chemistry 105 (2001) 11363.
[13] C. Brattwall, P. Lincoln, B. Nordén, Journal of the American Chemical Society 125 (2003) 14214.
[14] M.A.R.B. Castanho, S. Lopes, Journal of Physical Chemistry B 106 (2002) 7278.
[15] L.B.-Å. Johansson, G. Lindblom, Quarterly Review of Biophysics 13 (1980) 63.
[16] D. Oesterhelt, W. Stoeckenius, Nature New Biology 233 (1971) 149.
[17] W. Stoeckenius, R.H. Lozier, R.A. Bohomolni, Biochimica et Biophysica Acta 505 (1979) 215.
[18] R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckmann, K.H. Downing, Journal of Molecular Biology 213 (1990) 899.
[19] H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, J.K. Lanyi, Journal of Molecular Biology 291 (1999) 899.
[20] M.P. Heyn, B. Borucki, H. Otto, Biochimica et Biophysica Acta 1460 (2000) 60.
[21] K. Fahmy, F. Siebert, M.F. Großjean, P. Tavan, Journal of Molecular Structure 214 (1989) 257.
[22] M.S. Braiman, L.J. Stern, B.H. Chao, H.G. Khorana, Journal of Biological Chemistry 262 (1987) 9271.
[23] B. Albinsson, M. Kubista, B. Nordén, E.W. Thulstrup, Journal of Physical Chemistry 93 (1989) 6646.
[24] J.P. Cartailler, H. Luecke, Annual Review of Biophysical and Biomolecular Structure 32 (2003) 285.
[25] W. Meijberg, P.J. Booth, Journal of Molecular Biology 319 (2002) 839.
[26] D. Oesterhelt, L. Schuhmann, FEBS Letters 44 (1974) 262.
[27] W.I. Gruszecki, M. Gagos, M. Herec, P. Kernen, Cellular and Molecular Biology Letters 8 (2003) 161.
[28] A.M. Cerdeño, M.J. Bibb, G.L. Challis, Chemistry and Biology (2001) 817.
[29] R.R.C. New (Ed.), Liposomes: A Practical Approach, Oxford University Press, Oxford, 1989.
[30] R. Sarges, B. Witkop, A.V. Gramicidin, Journal of the American Chemical Society 87 (1965) 2011.
[31] T. Galbraith, B.A. Wallace, Faraday Discussions 111 (1998) 159.
[32] R. Marrington, T.R. Dafforn, D.J. Halsall, M. Hicks, A. Rodger, Validation of new micro-volume Couette flow linear dichroism cells, Analyst 130 (2005) 1608.

DOI: 10.1002/chem.200501168

Simulations of DNA Coiling around a Synthetic Supramolecular Cylinder That Binds in the DNA Major Groove

Syma Khalid,[a, c] Michael J. Hannon,[a, b] Alison Rodger,[a] and P. Mark Rodger*[a]

Abstract: In this work we present the results of a molecular simulation study of the interaction between a tetracationic bis iron(II) supramolecular cylinder, $[Fe_2(C_{25}H_{20}N_4)_3]^{4+}$, and DNA. This supramolecular cylinder has been shown to bind in the major groove of DNA and to induce dramatic coiling of the DNA. The simulations have been designed to elucidate the interactions that lead the cylinder to target the major groove and that drive the subsequent DNA conformational changes. Three sets of multi-nanosecond simulations have been performed: one of the uncomplexed d(CCCCCTTTTTCC)·d(GGAAAAAGGGGG) dodecamer; one of this DNA complexed with the cylinder molecule; and one of this DNA complexed with a neutralised version of the cylinder. Coiling of the DNA was observed in the DNA–cylinder simulations, giving insight into the molecular level nature of the supramolecular coiling observed experimentally. The cylinder charge was found not to be essential for the DNA coiling, which implies that the DNA response is moderated by the short range interactions that define the molecular shape. Cylinder charge did, however, affect the integrity of the DNA duplex, to the extent that, under some circumstances, the tetracationic cylinder induced defects in the DNA base pairing at locations adjacent to the cylinder binding site.

Keywords: bio-supramolecular chemistry · DNA recognition · helical structures · molecular dynamics · noncovalent interactions

Introduction

Control of gene expression is currently one of the key areas of interest in molecular medicine.[1–4] Gene expression involves the transfer of information encoded within the gene to produce a biologically active protein. However, not all genes are expressed in every cell all the time; bio-regulation of gene expression is effected by proteins that activate or repress transcription by binding to short, specific DNA sequences.[1,5] The ability to turn genes on or off artificially by the action of synthetic analogues of DNA-binding proteins is an important goal that would open up new possibilities for disease control and prevention as well as cure.

Proteins that bind DNA frequently achieve sequence-specific code recognition by binding non-covalently in the major groove of DNA.[6–8] In biological systems, the major groove is the preferred binding site for sequence recognition as this groove shows the greater variation in size and shape with base sequence and also the greater number and variation in pattern of hydrogen-bond donor and acceptor units; it is therefore the ideal target for synthetic molecules designed to recognise and bind to specific DNA sequences. However, relatively little progress had been made in producing synthetic major groove binders because of the size entailed. Most of the compounds synthesised are either minor groove binders,[9] intercalators,[10,11] or metal complexes that span only two to three base pairs and whose interaction with DNA is too limited for sequence selectivity.[12–14] Design of synthetic agents that target the major groove in a non-covalent, sequence-selective fashion remains a challenge of considerable importance in molecular medicine and biology.

[a] Dr. S. Khalid, Prof. M. J. Hannon, Prof. A. Rodger, Prof. P. M. Rodger
Department of Chemistry, University of Warwick
Coventry CV4 7AL (UK)
E-mail: p.m.rodger@warwick.ac.uk

[b] Prof. M. J. Hannon
Current address:
School of Chemistry, University of Birmingham
Birmingham B15 2TT (UK)

[c] Dr. S. Khalid
Current address:
Department of Biochemistry, University of Oxford
South Parks Rd, Oxford, OX1 3QU (UK)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

Recently, a major step has been taken toward achieving the goal of designing synthetic agents to target the major groove: we have developed a novel compound that binds strongly in the major groove of DNA and is large enough to span five base pairs.[15,16] The compound is a metallo-supramolecular tetracationic cylinder (**1**) with a triple helical framework. This cylinder is approximately 19 Å in length and 11 Å in diameter. It is too big to bind in the minor groove of DNA, but has just the right shape and size to lie along the major groove.[15]

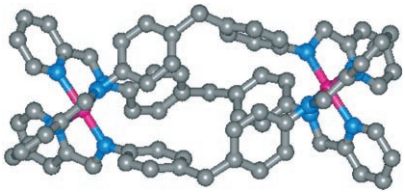


Figure 1. Tetracationic cylinder $[Fe_2L_3]^{4+}$, where L is *N*,*N*′-bis(pyridin-2-ylmethylene)-4,4′-diaminodiphenylmethane: Fe^{II} is pink, N is blue, C is grey; H atoms have been omitted for clarity.

Experimental techniques have provided information regarding the binding strength and preferred binding sites of **1** on DNA.[16] Spectroscopic studies indicate that the cylinder binds very strongly to DNA with a binding constant well in excess of $10^7\,\text{M}^{-1}$ in 20 mM salt. Flow linear dichroism (LD) reveals coiling or bending of the DNA on addition of the cylinder, and tapping mode atomic force microscopy (AFM)[16] reveals this to be dramatic intramolecular coiling of DNA that is unprecedented with a synthetic agent: **1** was found to induce DNA to bend by 40–60° per ligand,[15] with the smaller angles (40°) occurring at the maximum loading of one cylinder per DNA turn. In contrast cobalt ammines—which are known to cause DNA to bend and ultimately to condense into multi-molecular aggregates—generate just 2–5° bend per ligand.[17] The effect of cobalt ammines can largely be explained by the Manning charge condensation theory,[18] but the magnitude of the effect with **1** suggests the bending is of a very different character. The *M*-helical enantiomer of the cylinder coils DNA more aggressively than the *P* enantiomer.[16] NMR data have confirmed a major groove-binding mode for the *M* cylinder.[19] From all the experimental data it is clear that the cylinder binds in the major groove and is able to induce dramatic conformational changes in the DNA. However, gaining molecular level information about the effects and interactions which lead to this remarkable supramolecular event is challenging. Molecular dynamics (MD) simulations can provide information at the molecular level that is complementary to experiment and therefore are an ideal way to get a better understanding of this system.

There is comparatively little modelling literature on the interaction between transition-metal complexes and DNA. Some modelling has been performed on DNA with ruthenium(II)–tris(1,10-phenanthroline)[20] in vacuum, and the results compared with experimental data.[12] The Λ-ruthenium complex showed a preference for the major groove of DNA but little or no bending of the DNA at low loading and whereas the Δ enantiomer preferred the minor groove at low loadings where it did bend the DNA—but only by a few degrees—and bound to both grooves at higher loadings with no additional bending. Several other papers have used modelling studies to interpret experimental data on DNA/transition-metal complexes,[17,21] but since their emphasis was on understanding the experimental data, these calculations again omitted the solvent. Rigorous molecular dynamics simulations with explicit solvent have been performed for cobalt(III)–hexammine[22] and a nickel(II)–metalloprotein[23] in the presence of DNA. In both cases, the simulations produced good agreement with a range of experimental data; in neither case was there significant bending or coiling of DNA.

In this paper we present results of a molecular dynamics (MD) study of the effect of the supramolecular cylinder, **1**, on DNA. Docking calculations, both manual and using high-temperature MD with constrained or rigid molecules, have been used to identify favourable initial binding sites for the cylinder on a DNA dodecamer, and then multi-nanosecond MD simulations performed with explicit solvent to monitor the DNA response. Previously, MD simulations[19] were performed on this system to interpret the NMR spectra, but these used NOE data as additional force-field restraints. The simulations described in the present paper were designed to probe any changes in DNA structure upon cylinder binding and are not restrained in any way. As a consequence, these new simulations enable a direct analysis of the DNA response—both its nature and the forces that give rise to it.

Computational Methods

System: The DNA dodecamer sequence (CCCCCTTTTTCC)·d(GGAAAAAGGGGG) was chosen for this study. Preliminary simulations were performed with a decamer, d(CCCCCTTTTT)·d(AAAAAGGGGG), which was chosen since it might offer potential for identifying preference of A–T versus G–C tracts within a single series of simulations; this sequence was subsequently extended by adding two C–G pairs to the A–T end so as to minimise any tendency for the DNA ends to fray in the absence of the cylinder; note that no restraints were introduced into the force field to hinder end-fraying. The numbering used in this paper to identify specific bases and base pairs is defined below.

5′	C1	C2	C3	C4	C5	T6	T7	T8	T9	T10	C11	C12	3′
3′	G24	G23	G22	G21	G20	A19	A18	A17	A16	A15	G14	G13	5′
base pair no.	1	2	3	4	5	6	7	8	9	10	11	12	

The *M* enantiomer of the cylinder, **1** (denoted C^{4+}), was modelled using the CHARMM22 all-atom force field, but with the FeN_6 sub-unit treated as a rigid body with geometry taken from the crystal structure. For comparison, a simulation was also performed in which the *M* cylinder was

made electrically neutral, denoted C^0. Although the C^0 system has no experimental counterpart, a comparison of the C^{4+} and C^0 systems enables an exploration of the relative importance of molecular shape (as defined by the van der Waals interactions) and electrostatic interactions in inducing the DNA response. Analogous comparisons with netropsin[24,25] have proved useful in understanding the influence of charge in minor-groove binding ligands. Many methods could be used to achieve this neutrality, including setting all atomic charges to zero, neutralising the Fe atoms, or adding a neutralising negative charge to the organic part of the cylinder. In this work we have adopted the last option: $-0.03\,e$ was added to the charge on each C and H atom in the cylinder, resulting in a counterbalancing $-4\,e$ charge spread over the surface of the cylinder. While still giving rise to electric field gradients, the negative surface charge should considerably reduce the *electrostatic* attraction of the cylinder to the DNA phosphate backbone, and thus makes the comparison with C^{4+} instructive.

Technical details: Compounds were modelled using the CHARMM force-field series. This contains two main variants for modelling DNA: the older CHARMM22 force field,[26] and a more refined CHARMM27.[27] There have been a number of comparisons of how well these two force fields reproduce the behaviour of DNA,[28] and the general consensus is that the CHARMM27 force field is much better. In particular, it correctly predicts B-DNA to be the stable conformation in low ionic strength solvents at normal temperatures, whereas the CHARMM22 force field tends to cause the DNA to adopt an A-like form. However, it should be noted that most of these comparisons have treated just the duplex in aqueous solution: no comparisons have been reported for DNA interacting with a major-groove-binding ligand, particularly when that ligand carries a substantial positive charge. The experimental data discussed above indicates that C^{4+} perturbs the DNA strongly, bending it well away from the canonical B-form. This is outside the parameter-space that has so far been used to derive and validate the DNA force fields, and so it is of interest to compare the performance of these two force fields in modelling DNA/C^{4+} complexes.

All simulations were performed with DL_POLY.[29] The conversion from CHARMM to DL_POLY force-field formats was achieved by using a purpose-built program that interprets the CHARMM prm and crd files. As in previous work[30] checks on the force field were performed on numerous configurations to ensure the energies and forces calculated with DL_POLY and CHARMM agreed *exactly*. Water was modelled with the TIP3P[31] potential, and kept rigid using the SHAKE algorithm,[32] implemented in DL_POLY with a tolerance of 0.0001. All hydrogen atoms were assigned a mass of 2 u. This gave good energy conservation in constant energy simulations with a 2 fs time step for the CHARMM22 force field. An Ewald sum[33] was used to evaluate electrostatic interactions, with $\boldsymbol{k}_{max} = (5,5,6)$ and $\alpha = 0.12604\ \text{Å}^{-1}$. Simulations were performed with orthorhombic periodic boundaries at constant volume and temperature (NVT) using the Hoover thermostat[33] with a time constant of 0.1 ps and a time step of 2 fs; some 0.25 ns segments of the trajectories were repeated with a 1 fs time step and showed no significant differences. Some difficulties were encountered initiating CHARMM27 simulations with the 2 fs time step, and so this was reduced to 1 fs for all the CHARMM27 simulations reported herein. In all other respects, the CHARMM22 and CHARMM27 simulations followed the same protocols and methodology.

Protocol: The multi-nanosecond simulations were performed using a four-step protocol designed to identify good binding sites while still ensuring a disperse background ion distribution:

1) High temperature MD simulations were performed on the cylinder and ions moving around an immobilised B-DNA dodecamer, with the relative dielectric constant $\varepsilon_r = 80$;
2) low potential energy configurations were identified and optimised with respect to just the cylinder position; ions were kept fixed at their positions from the high temperature snapshot and the DNA also remained fixed;
3) water was added, the relative dielectric constant reset to 1, and the system allowed to relax while restraining the DNA with a series of progressively weaker harmonic tethering potentials;
4) full, unrestrained MD simulations were performed on the aqueous system for 2–5 ns.

This protocol is described more fully hereafter.

Docking (steps 1 and 2): A dodecamer of B-DNA and one cylinder molecule (*M* enantiomer) were embedded in a neutralising atmosphere of Na^+ ions. A further 58 Na^+ and 58 Cl^- ions were added in a $45\ \text{Å} \times 45\ \text{Å} \times 60\ \text{Å}$ box, with DNA aligned along the z (long) axis; this gave [NaCl] = 0.8 M, which is similar to the concentration used in some other MD studies.[34,35] The DNA was immobilized and an NVT MD simulation performed for 1 ns at 900 K; this temperature was found to be high enough to ensure that the ligand sampled the whole surface of the DNA efficiently, but without any significant deformation in its own shape. The temperature was also high enough to ensure that the background atmosphere of sodium and chloride ions behaved like a homogeneous gas, with no aggregation of the ions onto the DNA or cylinder. The conformations with the lowest configurational energy were extracted and energy-minimised with respect to the position of just the ligand. No change was allowed in the position of the DNA atoms, Na^+ ions or Cl^- ions during this optimisation of the cylinder within its DNA/ionic gas environment. The resulting configuration was then used to start fully solvated MD simulations as described below. Variants on this procedure were tried in which the DNA and/or Na^+ and Cl^- ions were harmonically restrained rather than frozen during the cylinder minimisation, but these produced essentially the same results when carried forward into fully solvated MD simulations. Manual docking calculations were also performed, but did not give more favourable docking sites than those identified using this high temperature MD docking protocol.

A number of alternative low energy conformations was examined and all found to involve major groove binding. Several of these were then carried forward into full solvated MD simulations following the protocol given below.[36] All such simulations exhibited a very similar DNA response to that reported herein—particularly with respect to the extent to which they bent the DNA, and the resultant stability of the DNA—and confirmed that the results of our simulations were not sensitive to variations in the initial binding site. We note that these repeat simulations also proved our results to be robust with respect to the initial background ion distribution, since each configuration also involved a very different, essentially random, initial arrangement of the sodium and chloride ions.

Molecular dynamics simulations (steps 3 and 4): Water molecules, taken from an equilibrated liquid water simulation, were added to the optimised configurations (DNA, cylinder, Na^+ and Cl^- positions) identified in the docking calculations. Any water molecule that overlapped with DNA, cylinder, Na^+ or Cl^- was removed, which left a total of 3720 water molecules in the final DNA/ligand system, or 3758 for the uncomplexed DNA. Equilibration followed a similar protocol to that used by other research groups.[23,28,37] The DNA atoms were tethered to their original positions with a harmonic force constant of $100\ \text{kcal mol}^{-1}\ \text{Å}^{-2}$ and an NVT MD simulation performed for 10 ps at 310 K. A further five 10 ps simulations were then performed successively in which the tethering force constant was 50, 25, 10, 5 and $1\ \text{kcal mol}^{-1}\ \text{Å}^{-2}$, respectively. The tethering potentials were then removed completely and multi-nanosecond simulations performed, saving configurations every 1 ps for later analysis.

Up to four repeat simulations of 2–5 ns duration, using slightly different starting configurations, were performed for each ligand/DNA system to validate the reproducibility and timescale of the DNA response. In every case, the response of the DNA was found to be rapid and repeatable, with most changes occurring in the first 0.5 ns, and no noticeable changes occurring between 2 and 5 ns.[36] In this paper we therefore concentrate on the behaviour during the first 2 ns after binding.

Analysis methods: A number of different methods have been used in this work to monitor DNA structure. The time dependence of conformational parameters has been monitored using the Curves algorithm[38] implemented in MDToolchest.[39] This uses seven torsion angles (α–ζ, ϕ, χ) to describe the DNA backbone, while the intra-base-pair geometry is described by six *helicoidal* parameters: three displacements (shear,

stretch and stagger) and three angles (buckle, propeller and opening). This parameter set is now well established as a means of describing DNA conformation;[40,41] a full definition is supplied with the Supporting Information, though it is useful to note here that of the backbone angles, α and ζ refer to torsions about P–O bonds, δ and ϕ refer to the ribose ring, while χ is for the bond that links the backbone to the base.

In understanding the DNA conformation and flexibility, it is also useful to examine some of the parameters that give a more correlated view of the overall behaviour. In particular, it is useful to study the relationship between neighbouring base pairs. In this work these have been monitored using a distance (slide) and three angles (tilt, roll and twist; see Figure 9) as implemented in 3DNA.[42] In this case, the analysis has been applied to the average structure determined from a continuous 50 ps portion of the MD trajectory; this length of trajectory was found to be long enough to smooth out the instantaneous fluctuations in shape, but was still short with respect to the systematic relaxation induced by the presence of the cylinder, **1**. The 3DNA analysis of average structures has also been used to generate normal vector plots,[7,43] which are useful for identifying linear and bent regions within the DNA, and hence to estimate the degree to which the DNA is bent by the cylinder.

Results and Discussion

Response of the DNA to a bound cylinder

As will be discussed below, the simulations with the two DNA force fields (CHARMM22 and CHARMM27) gave broadly similar responses for the DNA/ligand complex, in that the binding site was consistent with the two models and both indicated the DNA to be bent by the cylinder. Some quantitative differences were observed between the two force fields, with CHARMM27 tending to give the more minimal response to the cylinder. Two examples (discussed in more detail below) are worthy note at this stage. The overall bend of the DNA was about 40° with CHARMM22 compared with 20° for CHARMM27, and the CHARMM22 DNA did respond differently to C^{4+} and C^{0} whereas the CHARMM27 model was remarkably insensitive to this large change in ligand charge. In the following we focus first on the results for the CHARMM22 model: since the response is greater with this model it is likely to be easier to identify the underlying physics of the DNA/ligand interaction. This is then followed by a more detailed comparison of the force fields to place confidence limits on the interpretation of the simulations.

Charmm22: As described in the protocol, favourable binding sites for the cylinder on the DNA were identified and then explicit solvent molecules added. In all cases, binding was found to occur preferentially in the region of the A–T tract (base pairs 6–10). Using the position from the docking calculation as a starting point, water was added, the system equilibrated while the DNA conformation was restrained to the B-form, and then simulations allowed to proceed without restraints. We re-iterate that up to four such simulations were performed for each DNA/cylinder system.

Once the restraints were removed, the DNA was observed to respond rapidly to the presence of the cylinder. Major conformational changes were observed in the presence of both C^{4+} and C^{0}. In every trajectory these occurred within the first 0.5 ns, with the precise timing depending on the starting configuration. In all cases, no significant conformational changes were observed in the subsequent 4–5 ns, suggesting that our simulations were probing the final DNA state, or at least identifying a long-lived intermediate in the response process.

It is instructive to begin by comparing before and after images of the DNA conformation in the three different systems: uncomplexed, bound to C^{4+}, and bound to C^{0}. Although such snapshots provide only limited information about the range of DNA conformations, for the present study they prove to give a useful and visual overview of the more quantitative data presented below. Images of the initial (directly after docking) and relaxed (after 2 ns) configurations for DNA and cylinder are given in Figure 2. Initially, the DNA adopts a nearly linear B-form. The initial binding mode of the cylinder is in the major groove, lying symmetrically between the two strands of the DNA, and spanning base pairs 6–11 (T6–A19 to C11–G14). This is consistent with the published structure obtained from NMR (NOE) data.[15,16,19] This binding geometry was retained throughout the equilibration phases, and was not disrupted until the tethering potentials that restrained the DNA conformation were removed.

After 2 ns of the unrestrained simulations, the *uncomplexed* DNA retained its near linear B-form, albeit with some fraying at the first base pair. In contrast, the final configurations of the DNA/cylinder simulations suggest substantial curvature of the DNA for both charged and uncharged cylinders. As will be discussed below (see Figure 7), these snapshots do depict a bend of about 40° in the helix axis in the presence of the cylinders.

At the end of the simulations the charged cylinder, C^{4+}, remained in the major groove, but some distortion of the groove shape was evident. There was close association between the cylinder and base-pairs 5–9 (closest interatomic distances between the cylinder and each of these nucleic acids are 2.4–3.5 Å); thus the cylinder remained within the A-T tract. There was no evidence of hydrogen bonding between the cylinder and DNA, but this is not surprising since the cylinder has no conventional hydrogen-bond donor groups, and its potential acceptors are obscured by nonpolar carbons. It is particularly interesting to note that some breakdown of the Watson–Crick structure is apparent in the last three or four base pairs of the DNA (Figure 2, middle, bottom), that is, adjacent to the cylinder binding site. As is shown below, this is due to a mispairing between T9 and A15, which then leaves T10 and A16 unpaired.

Importantly, the simulations reproduce both the major groove binding characterized by NMR and the DNA coiling effects as observed in LD and AFM experiments. The implication of the simulations is that, at least for some sequences, this bending might be associated with alterations to the DNA base pairing. Such effects have also been observed in some DNA–protein complexes,[44,45] and can, indeed, be an important feature of the way that proteins process DNA.

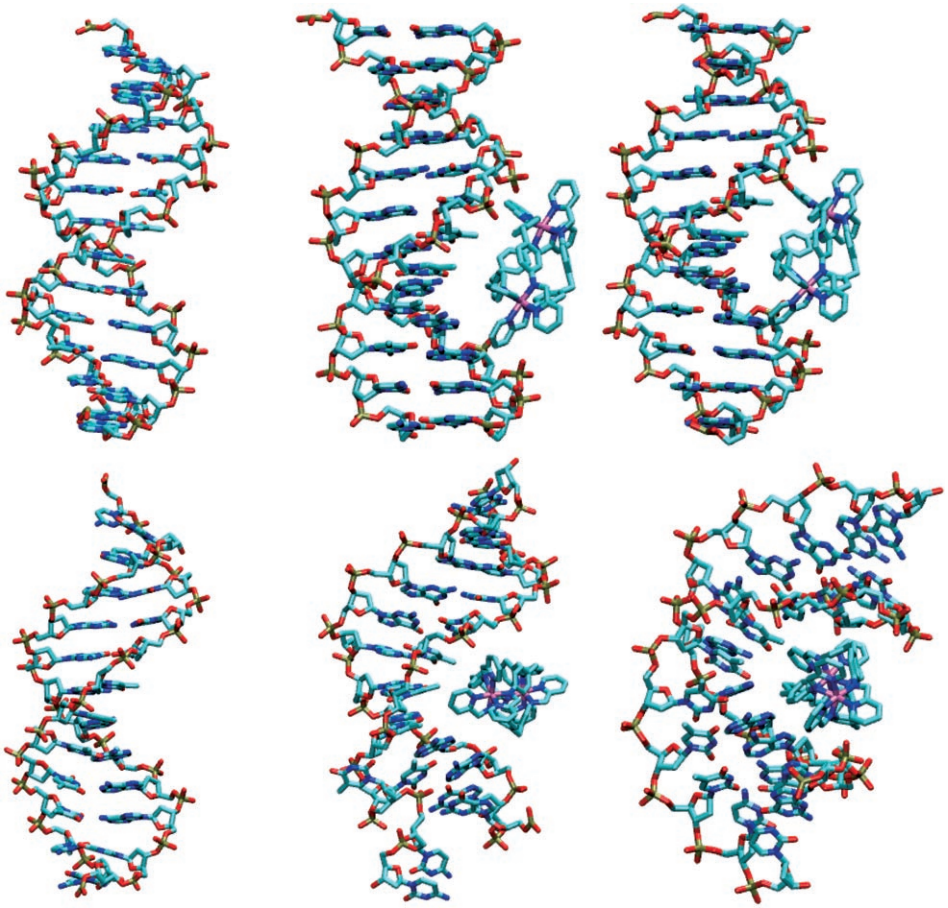


Figure 2. Conformations of DNA taken from the beginning (top) and after 2 ns (bottom) of MD simulations using the CHARMM22 force field: uncomplexed DNA (left), DNA/C^{4+} (middle) and DNA/C^{0} (right). Conformations after 5 ns were very similar to those at 2 ns.

Moreover, DNA bending is commonly associated with A–T tracts.[46,47]

The behaviour of the neutralised cylinder, C^{0}, provides a fascinating contrast to that of C^{4+}. Its final configuration shows C^{0} to lie symmetrically within the major groove, spanning base pairs 5–10, but with the DNA having bent substantially to create a much deeper pocket for the cylinder than was found with C^{4+}. It might originally be thought that the coiling of the DNA was, in large part, due to the attraction between the large positive charge on C^{4+} and the negatively charged phosphates in the DNA backbone. However, it is clear from Figure 2 (and the quantitative analysis presented below) that the neutral cylinder, C^{0}, also causes strong coiling in the DNA, but does so without inducing any of the disruption of the Watson–Crick base pairs that was evident with C^{4+}. This contrast suggests that the shorter ranged van der Waals forces play a significant, perhaps even dominant, role in coiling the DNA backbone while the large Coulombic forces generated by the exposed cationic charge of C^{4+} may generate a localised stress that is strong enough to disrupt inter-strand hydrogen bonding in the DNA duplex.

The helicoidal parameters developed by Lavery and Sklenar[38] are a good source of quantitative data on how these ligands affect the integrity of the base pairs within the DNA double helix. These parameters describe how two nucleic acid bases move and rotate relative to each other and are defined such that zero displacement/rotation corresponds to the ideal base pair geometry for B-DNA. Helicoidal parameters have been calculated from all three simulations and are presented in Figure 3. For the uncomplexed DNA, nearly all the base pairs can be seen to be stable, with typically small fluctuations about values of zero (the ideal base-pair geometry). The only exception is the first pair, C1–G24, which, from about 1 ns, showed large amplitude motions characteristic of fraying at the end of the double helix; end-fraying is a not uncommon event for DNA, both in simulations and in vivo. Some large-amplitude fluctuations were also seen for T8–A17 and T9–A16 pair at about 1 ns, but these were transient and rapidly returned to stable values; such behaviour is indicative of the flexibility inherent within the DNA duplex rather than of irreversible conformational changes.

Given the amount of curvature evident in Figure 2 (right), the C^{0} ligand gives rise to helicoidal parameters that are remarkably similar to those of the uncomplexed DNA. Indeed, the principal difference between these two sets of data is that the transient large amplitude oscillations seen in the uncomplexed DNA are completely absent for DNA in the presence of C^{0}. As with the uncomplexed DNA, fraying is again found only at the first base pair. Thus, the presence of C^{0} actually appears to enhance the stability of the Watson–Crick base pairs, despite the overall curvature of the DNA evident in Figure 2 (right).

For C^{4+} bound to DNA, the helicoidal parameters clearly reveal a deformation of the base-pair stack. There are two distinct zones in the double helix for this system. The first seven base pairs show stable behaviour as seen in the uncomplexed DNA, with just small fluctuations around zero. Similar behaviour is seen for T8–A17, albeit with larger amplitude vibrations for the shear (SHR), buckle (BKL) and open (OPN) parameters. However, the plots for the last four base pairs show extremely large variations that are

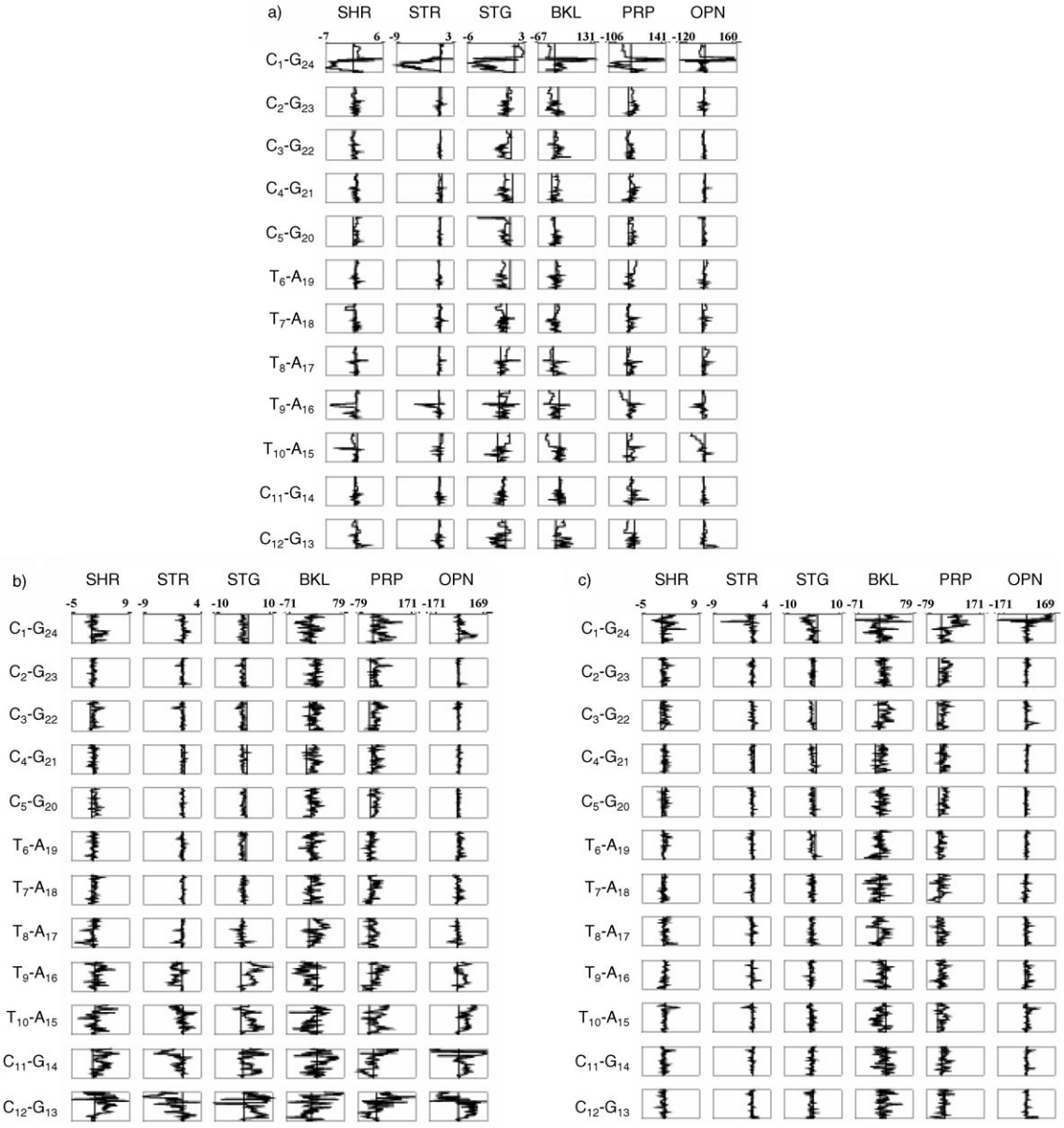


Figure 3. Helicoidal parameters for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^{0}. Time is given on the vertical axis (0–2 ns, 0 top) and the helicoidal parameter on the horizontal axis (scale as marked). For a definition of these parameters[38] see Supporting Information.

simply inconsistent with a stable Watson–Crick base pair: shear deformations vary by as much as 8 Å during a simulation, while the base pair "stretches" are, at times, 3–4 Å *less* than their equilibrium value.

Visual inspection of the DNA/C^{4+} trajectory showed that this apparent disruption of the double helix below base-pair T9–A16 was actually a complex rearrangement amongst the base pairs, resulting in the formation of a new pairing between T9 and A15, so that T10 and A16 were left unpaired (see Figure 4). The sequence of events for the formation of this defect was as follows:

1) Some stress became evident in base pairs 10–12 from about 170 ps:
2) T9–A16 broke apart at 390 ps;
3) a new base pair, T9–A15, formed within the following 10 ps.

T10 was then observed to swing out of the helix, kinking the backbone, and subsequently disrupting the last two base pairs (C11–G14 and C12–G13). It is interesting to note that this activity occurred adjacent to, rather than at, the cylinder binding site: the cylinder remained bound to the base pairs numbered 5–8 (i.e., C5–G20 to T8–A17) throughout the simulation.

This change observed in the base-pair stacking is also revealed in the pattern of hydrogen bonding between the nucleic acids at the end of the simulation. Hydrogen-bond

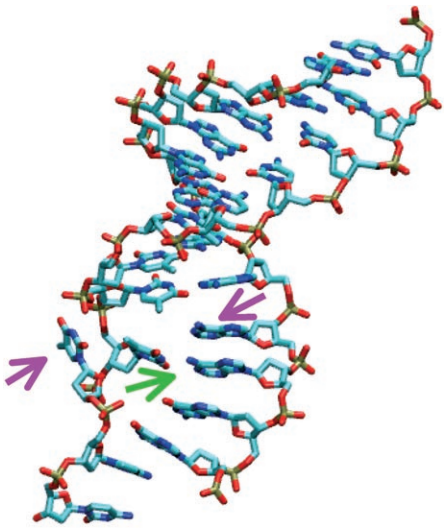


Figure 4. Snapshot of DNA showing the base-pair defect induced by C^{4+}. The green arrow points to the slipped mismatch between T9–A15; the magenta arrows indicate the resulting unmatched bases, T10 and A16.

lengths were obtained from the average DNA structure, as calculated by the 3DNA package, using the final 50 ps of each trajectory. This time window was found to be long enough to smooth out the instantaneous vibrations, but still short enough to avoid artefacts due to long timescale conformational motions of the DNA. The results for the two DNA/cylinder simulations are listed in Table 1 and confirm the formation of a mismatched Watson–Crick base pair T9–A15, with the two hydrogen bonds, in the presence of C^{4+}.

Dial plots of the DNA backbone torsion angles from the three sets of simulations are given in Figure 5. Data for the uncomplexed DNA are consistent with the average values expected for B-DNA, and show no evidence of a transition to one of the other forms. The presence of the cylinder does change the DNA backbone conformation, but the changes are smaller than was seen in the helicoidal parameters. Interestingly, while the helicoidal parameters indicated that the base-pair structure was more constrained in DNA/C^0 than in the uncomplexed DNA, the opposite appears to be true for the backbone, with the backbone torsions showing larger fluctuations in the presence of C^0 than without. The changes in base-pair stacking induced by C^{4+} are also seen in the backbone, but the difference between the two cylinders is less obvious than it was with the direct measures of base-pair geometry. The most obvious indication is in the C3′-O-P-O (ζ) and O-P-O-C5′ (α) angles, which show almost random angular variations between base pairs 10–12 in the presence of C^{4+}, suggestive of rotation that is unhindered by base-pair formation. This effect appears to be more pronounced at the end of the first strand (nucleotides T10 to C12) than on the complementary strand (nucleotides G13 to A15).

A schematic representation of the global helix axis after 2 ns is given in Figure 6. We note that quantifying curvature in the DNA helix can be problematical. Any measure must be able to distinguish between the local oscillations in base-pair orientation found in, particularly, the A-form, and a real bend in the overall helix direction. This can be particularly difficult in short strands of DNA, such as the dodecamer studied here, since one cannot use the behaviour of the helix beyond the binding site to confirm the persistence of any bend. To measure the extent of bending we have used the 3DNA definition of the local helix direction. This uses the geometry of any two adjacent base pairs (a step) to define a unit vector, $\boldsymbol{U}_i$, along the helix axis at that step. Bending of the DNA can then be monitored by defining an angle, $\vartheta_i = \cos^{-1}(\boldsymbol{U}_{ref} \cdot \boldsymbol{U}_i)$, which describes the total bend in the DNA between some reference step (taken to be the first stable step in this work) and the *i*th base-pair step. Most importantly, this definition gives perfectly aligned $\boldsymbol{U}_i$ ($\vartheta_i = 0$ for every *i*) in *both* the canonical A and B forms, so that deviations from 0 can be interpreted as bending of the DNA.

ϑ_i values for the, averaged, DNA structures after 2 ns are shown in Figure 7; we reiterate that the DNA response was complete on this timescale, and no further relaxation was observed in simulations continued to 5 ns. Thus, these plots give a good indication of both the global bend induced in the DNA by the ligands, and how it is distributed along the dodecamer. Both C^{4+} and C^0 can be seen to bend the DNA by about 40°, which compares very favourably with the 40–60° per cylinder measured experimentally. In the presence of C^{4+} this bend is strongly localised on steps 5 and 6 (base pairs 5–7), but with the consequent degradation of base pairs 9–12 noted above. In contrast, the

Table 1. Intra-base pair hydrogen bond length, taken from the average DNA structure observed during the final 50 ps of the simulation; blank values indicate no hydrogen bond was found. The mismatch induced by C^{4+} is indicated in bold.

Base pairings	Length [Å]					
	DNA/C^{4+}			DNA/C^0		
	O2–N2	N3–N1	N4–O6	O2–N2	N3–N1	N4–O6
C1–G24	2.75	3.06	3.22			
C2–G23	2.96	3.04	2.97	2.82	3.00	3.02
C3–G22	2.86	3.00	3.01	2.86	3.01	2.97
C4–G21	2.80	3.03	3.10	2.80	3.15	3.35
C5–G20	2.91	3.05	2.99	3.03	2.99	2.84
T6–A19		2.98	3.18		2.99	3.01
T7–A18		3.05	2.91		3.06	2.95
T8–A17		3.03	2.89		3.00	2.87
T9–A16					3.03	3.00
T9–A15		**3.17**	**2.89**			
T10–A15					3.13	2.88
C11–G14				3.00	3.01	2.94
C12–G13				2.85	3.01	3.00

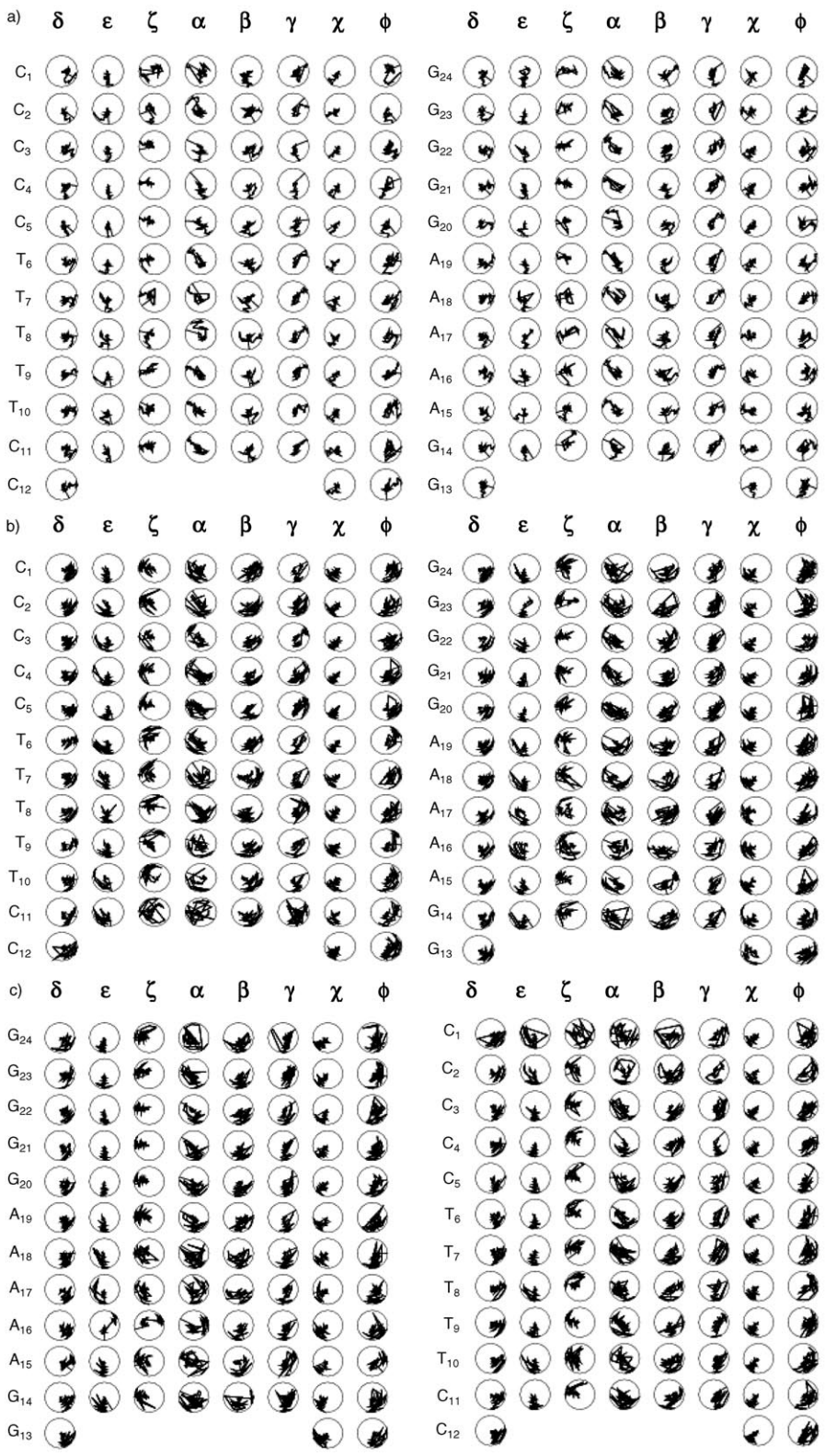


Figure 5. DNA backbone torsion angles for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^{0}. Circular polar plot of time (radius, 0–2 ns) against torsion angle (0° top, 90° right).

40° bend for C^{0} is spread over 10 base pairs, which probably explains why the integrity of the duplex is retained in this case. The uncomplexed DNA shows a net bend of only 10–12° over 10 base pairs, with some evidence of more pronounced curvature over the first few steps.

Another common descriptor for helix curvature is the normal vector plot (NVPs).[43] NVPs are projections of the unit vector normal to the plane of each base pair onto a plane that is normal to the helix axis. Linear segments of the DNA can then be identified as clusters of neighbouring points, while coiling is seen as a steady change of orientation across a number of adjacent base pairs. NVPs calculated at the end of each simulation are presented in Figure 8. The uncomplexed DNA shows a reasonably tight cluster of points. In the presence of the neutralised cylinder (C^{0}), however, there is a large and systematic variation in orientation across base pairs 3–9, consistent with the strong curvature noted above. For the fully charged cylinder, C^{4+}, there is strong curvature but as was seen with the cumulative bending angle (Figure 7), this effect is localised between base pairs 6 and 7, that is, to the start of the region in which the ligand is bound, and is suggestive of a localised kink in the double helix.

Various parameters describing the geometry of base-pair steps have also been calculated using 3DNA, and the average structure obtained during the final 50 ps portion of the trajectory. In Figure 9 we report values of the slide displacement, and roll, twist and tilt angles,[40,41] for all stable base pairs. There are several important points to note from these plots. In the first place, the step parameters observed in the presence of the cylinders are not readily identified with any of the common DNA conformations. Slide values below −1 Å, as observed along most of the stable duplex in the presence of both C^{4+} and C^{0}, are often taken to be indica-

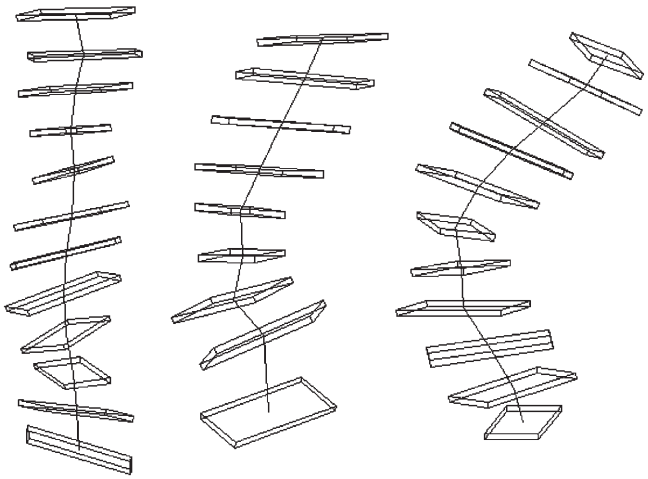


Figure 6. Representation of the average DNA structure in the uncomplexed (left), DNA/C^{4+} (middle) and DNA/C^0 (right) simulations. Averages were calculated from the final 50 ps of each simulation using 3DNA.[42] Only the stable base pairs are depicted.

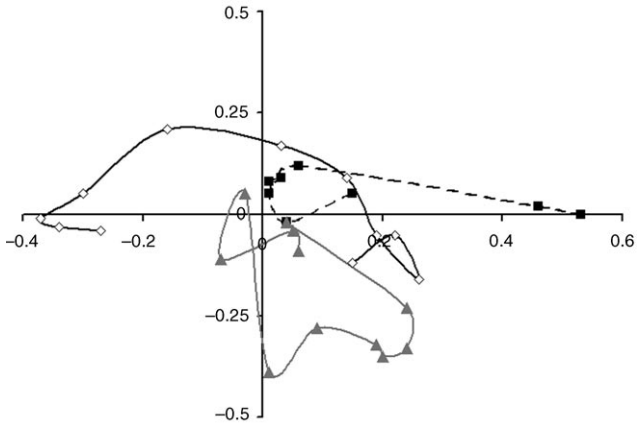


Figure 8. Normal vector plots (NVPs) for the average DNA structure calculated from the final 50 ps of each simulation using the CHARMM22 force field. NVPs are a projection of a unit vector normal to each base pair onto a plane that is normal to the local helical axis, with the points corresponding to successive base pairs being joined by a smooth line. ◇: C^0, ■: C^{4+}, ▲: uncomplexed.

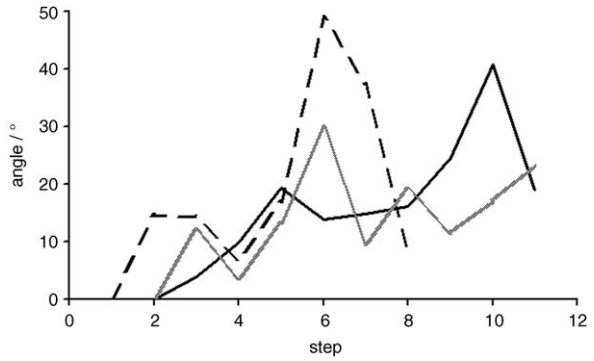


Figure 7. Degree of bend in the DNA, calculated from the local helix direction as defined by 3DNA. The angles are defined as $\vartheta_i = \cos^{-1}(\boldsymbol{U}_{ref} \cdot \boldsymbol{U}_i)$, where $\boldsymbol{U}_{ref}$ is the helix director for the first stable step. The helix bend (left) is identically zero for both the canonical A- and B-forms. ——: C^0, -----: C^{4+}, ——: uncomplexed.

tive of A-DNA. However, both the slide and roll values obtained in this work are really intermediate between the values expected for A- and B-DNA, while, for C^0, the twist is more consistent with B-DNA. On balance, 3DNA assigns the duplex as having the B-form in the presence of the neutral cylinder, and fails to make any assignment in the presence of C^{4+}. It is also interesting to note that both the slide and the roll angles take their most extreme values at step 6, which is both the start of the region to which the cylinders bound and the start of the A-T tract. Finally, we note that the twist angles indicate that neither C^{4+} nor C^0 cause significant winding or unwinding of the DNA—either locally or globally. The average twist for C^0 is 33.4° per step, which is intermediate between the A- and B-forms. A slightly lower value (29.4° per step) is seen in the presence of C^{4+}, but the extent to which this is influenced by the mismatched base pairing is unclear. This behaviour is in contrast to some proteins, such as 434 Cro, which bend the DNA via alternating regions of under- and overwinding.[1]

A concomitant question when DNA bending is observed is to ask how the size of major and minor grooves is affected. Following Hassan and Calladine,[48] we have monitored these by measuring the distance between appropriately displaced phosphate groups, with the P atom representing the position of the phosphate. The width of the major groove can be defined as the distance between the P atom in the *n*th nucleotide, counting from the 5′ end of one strand, and the P atom of the *n*+3rd nucleotide from the 3′ end of the complementary strand; this numbering assumes 5′ nucleotides. Similarly, for the minor groove one can use the distance between the P atoms in the *n*th nucleotide from the 5′-end of one strand with the *n*−4th nucleotide from the 3′-end of the complementary strand. The distribution of these inter-phosphate distances was determined by analyzing every configuration saved from the final 0.25 ns of each simulation. To avoid artefacts from end-fraying, those phosphates that were adjacent to the first and last base pair were omitted from the analysis. This left five measurements of the major groove width, and six for the minor groove. It should be noted that, for a dodecamer, this meant that the minor and major groove measurements were made on largely different regions of the DNA, with only two base pairs (T6–A19 and T7–A18) being used to measure both the major and the minor groove widths.

Unlike crystal structure analyses, the DNA shows considerable flexibility during a simulation, and so gives rise to a distribution of groove widths for each pair of phosphate groups. The average groove width and its standard deviation are presented in Figure 10. It is striking that the DNA-coiling induced by C^0 does not in any way affect the width of the major groove—the site into which C^0 binds. An enlargement in the width of the minor groove (ca. 2.5 Å) is seen in the vicinity of the C^0 binding site, but only for four of the

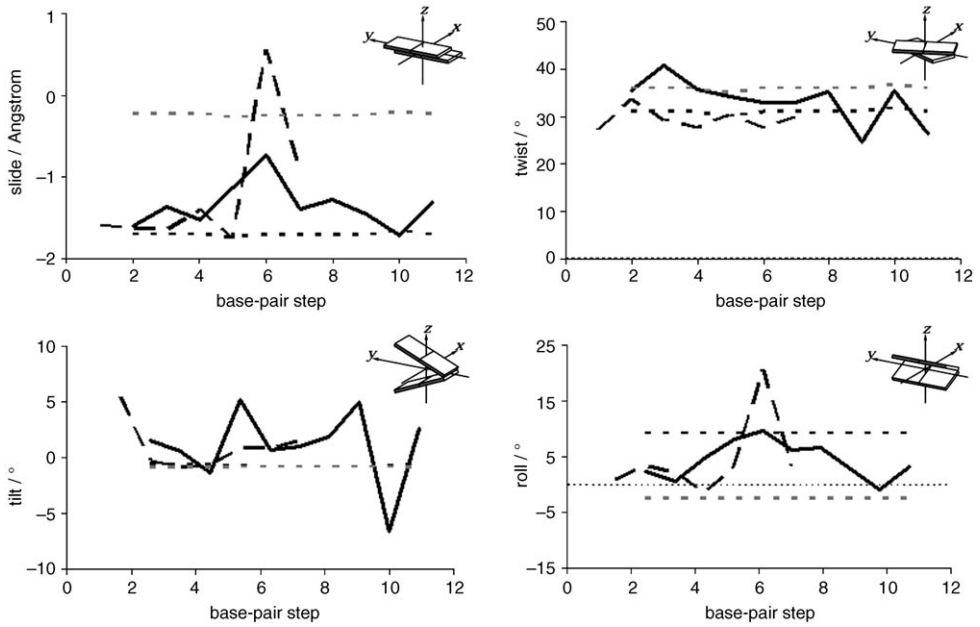


Figure 9. Slide displacement, and twist, tilt and roll angles between adjacent base pairs; a schematic definition of each inter-base pair coordinate is given with the plot. The horizontal axis corresponds to the step between base pairs, so that step 1 compares the first two base pairs (C1–G24 and C2–G23), etc. Only stable base-pairs have been included. Data for the canonical A- (-----) and B-forms (- - - - -) are given as a reference; ——: C^0, – – –: C^{4+}.

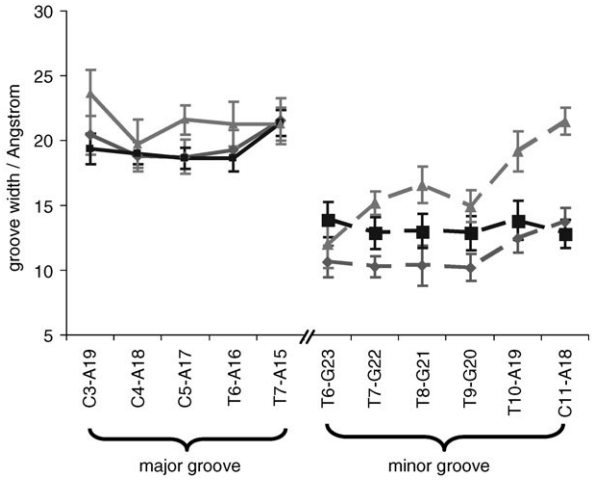


Figure 10. Widths of the major (solid) and minor (dashed) grooves, as determined from specified inter-phosphate distances. Calculations were performed with CHARMM22. "Error bars" denote the width of the distribution of distances (± one standard deviation), not the uncertainty in the calculation. No correction has been made for the size of the phosphate group; ◆: uncomplexed, ■: C^0, ▲: C^{4+}.

six inter-phosphate distances found in this region (those involving nucleosides T6 through T9, but not A18 and A19).

In contrast to C^0, C^{4+} has a substantial effect on the widths of both the major and minor groove. The major groove width increases by an average of 1.8 Å compared with both the uncomplexed DNA and the DNA with bound C^0. Even larger increases are seen in the width of the minor groove. However the largest increases (nearly 8 Å for C11–A18) are in regions where the backbone torsion angles α and ζ have already been shown to be ill defined (Figure 5) and so probably indicate that there is no well defined minor groove in this region when C^{4+} is bound to the DNA.

In summary, the supramolecular cylinder, **1**, is seen to have a strong influence on the conformation of DNA as modelled with the CHARMM22 force field. The interactions within a neutralised analogue of **1** were sufficient to induce a bend of about 40° in the DNA helix axis, which is very similar to the angle seen in experiments with **1**. When the tetracationic charge is also accounted for, **1** was found to strain the base pairing within the duplex, and induced a mispairing of A–T bases adjacent to the ligand binding site. It is important to stress that the induction of base-pair defects was reproducible. Four separate C^{4+} simulations were performed, starting with different low-energy cylinder/DNA binding sites and with different arrangements of the solvent and counterions. In each case a mismatched A–T base pair formed within the first 0.5 ns. It is possible that the defects were introduced as part of the coiling process (i.e., whether the additional strain during coiling made the DNA susceptible to defects in the base-pair sequence) in which case they would not reflect a global free energy minimum. To test this, a further simulation was performed starting from a stable configuration obtained after 2 ns of the DNA/C^0 simulation and changing the atomic charges back to those of C^{4+}; four Na^+ ions were also removed to maintain electrical neutrality in the system. An equivalent extended simulation with C^0 was used as a control. Within just 0.25 ns, mispairings in the A–T region again appeared in the presence of C^{4+} (Supporting Information, Figure 4), but no mispairings appeared with

C^0 during an additional 3 ns. We conclude that strong coiling of DNA by **1**, coupled with a disruption of the duplex base pairing, is a real prediction from the CHARMM22 force field.

Force-field dependence: The two DNA force fields have been compared by measuring the extent to which a stable binding geometry obtained with one force field is maintained by the other. This amounts to determining whether an important minimum on the free energy surface of one force field is reproduced with the other. By starting in a well-defined free-energy minimum, instead of repeating the docking protocol, one avoids the possibility of simulations with different potentials simply converging on different local minima and thereby ends up with an unambiguous comparison of the two force fields. Accordingly, simulations with the CHARMM27 force field were started from the stable conformation obtained after 2 ns in the C^0/CHARMM22 simulation.

Simulations were again performed with both tetracationic and neutralised cylinders and are denoted ${}_{27}C^{4+}$ and ${}_{27}C^0$, respectively. For the ${}_{27}C^{4+}$ system, four Na^+ ions chosen at random were removed to maintain electrical neutrality. The system was then relaxed in a series of short, restrained simulations, as outlined in Section Computational Methods. All restraints (i.e., tethering potentials) were then removed and the system simulated for an additional 1.5 ns. The 3DNA analysis of bend and step parameters, obtained from the average structure at the end of these simulations, is presented in Figures 10–12, while the distribution of groove widths observed during the final 0.25 ns is depicted in Figure 14. Where clarity permits, results for CHARMM22 are also depicted for comparison.

In general, the CHARMM27 DNA is less strongly influenced by the cylinder than is the CHARMM22 model. For both ${}_{27}C^{4+}$ and ${}_{27}C^0$, the DNA uncoils slightly to give an overall bend of only about 20° (Figure 11), while the normal vector plots (Figure 12) remain relatively compact. At the same time the slide parameters (Figure 13) indicate a shift to the B-form that is known to be favoured by CHARMM27 for an aqueous duplex in the absence of a ligand. No evidence was found that either cylinder disrupted the base pairings at any stage during the simulations, with the average structure from the end of the simulations showing the expected three (G–C) and two (A–T) hydrogen bonds for every base pair; at least one of the hydrogen-bond lengths was less than 3.0 Å in every base pair, while none was more than 3.13 Å. This is a major difference from the analogous CHARMM22 simulation described above, where reinstating the $+4\,e$ cylinder charge on the DNA/C^0 complex induced base-pair mismatches within just 0.25 ns.

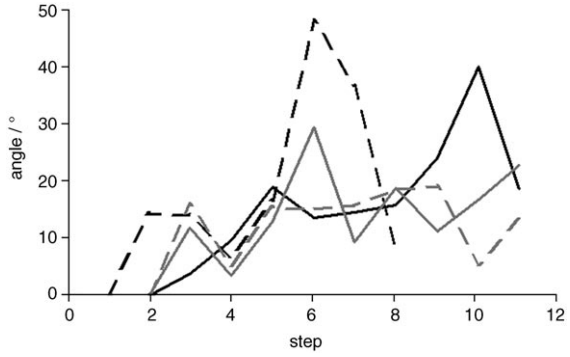


Figure 11. Global bend of the DNA, calculated as for Figure 6. CHARMM27 results are given as dashed lines, and CHARMM22 as the solid lines; ——: C^0, -----: C^{4+}, ——: ${}_{27}C^0$, - - - - -: ${}_{27}C^{4+}$.

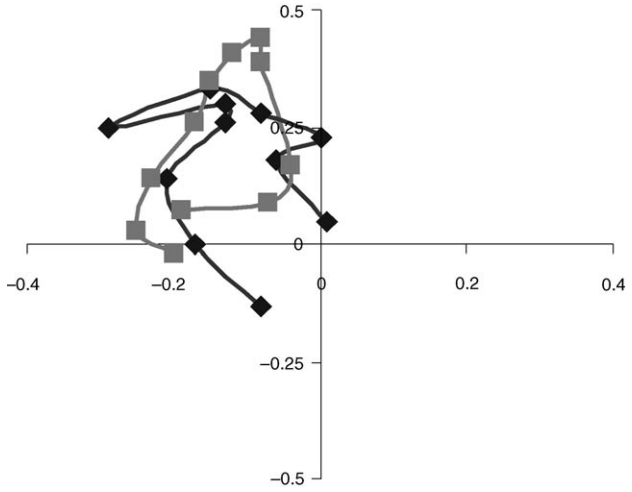


Figure 12. Normal vector plots for the average DNA structure obtained during the final 50 ps of the CHARMM27 simulations. ◆: ${}_{27}C^0$, ■: ${}_{27}C^{4+}$.

Perhaps most surprisingly, the cylinder charge appears to have very little effect on the CHARMM27 DNA. The normal vector plots, global bend and step parameters (with the exception of the roll angle at the 6th step) all show remarkably similar behaviour for the ${}_{27}C^{4+}$ and ${}_{27}C^0$ systems. Given the strongly ionic character of DNA and its flexibility, this apparent invariance to the ligand charge is unexpected. One property that does show differences between the two force fields is the groove widths. Relative to CHARMM22/C^0, the major groove of CHARMM27-DNA contracts in the presence of C^{4+}, but expands in the presence of C^0. We note that CHARMM27 uses a more strongly charged DNA backbone than does CHARMM22, with the PO_4 net charge being $-1.2\,e$, compared with $-1.0\,e$ for CHARMM22. This increased negative charge may lead the DNA backbone to contract more strongly onto the tetracationic cylinder, resulting in a stiffening and straightening of the DNA backbone and thereby inhibiting the subsequent bending of the DNA. A similar response to ligand charge has been seen with netropsin bound in the minor groove,[24] but greater scope for DNA relaxation might have been anticipated when the ligand binds in the major groove.

The comparison between these two force fields highlights several influences on the DNA response. In the first place we note that both force fields predict the neutralised ligand to coil DNA at least as strongly as the tetra cationic ligand. Previous work has shown that regions of low dielectric con-

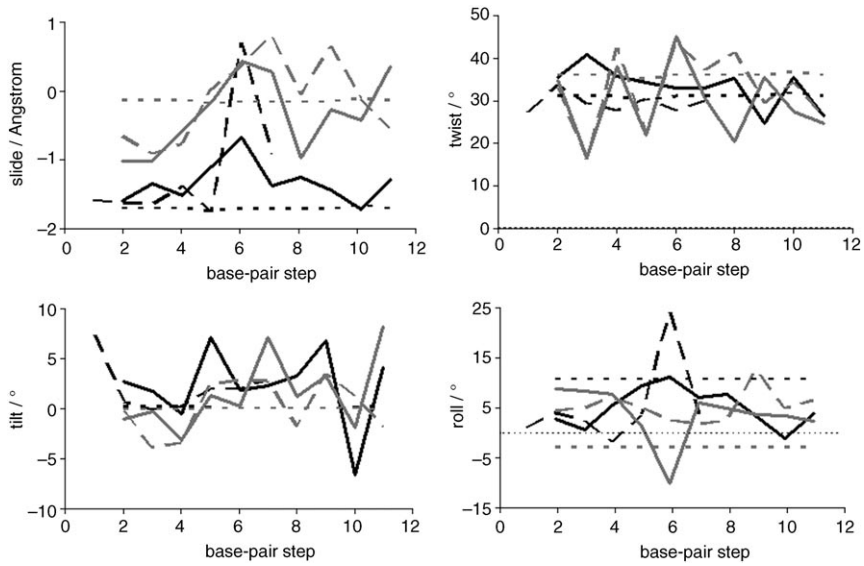


Figure 13. Selected step parameters for stable base pairs, calculated as for Figure 8. Values for the canonical A- and B-forms are given for reference, ——: C^0, -----: C^{4+}, ——: $_{27}C^0$, ------: $_{27}C^{4+}$, •••••: A-form, ·····: B-form.

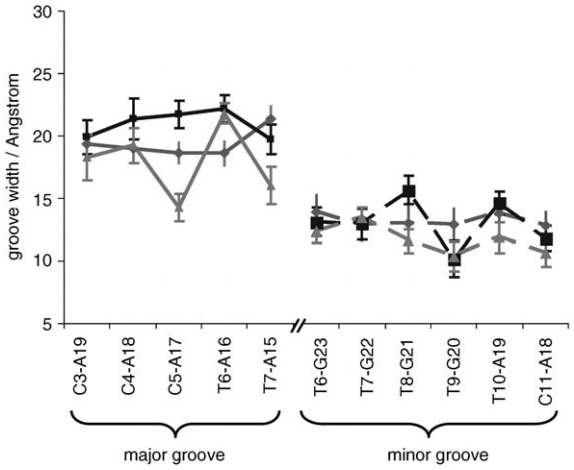


Figure 14. Groove widths calculated during the final 0.25 ns of the CHARMM27 simulations. Solid lines indicate the major groove and dashed lines the minor groove,◆: C^0, ■:$_{27}C^0$, ▲: $_{27}C^{4+}$.

stant enhance the phosphate repulsions in DNA and cause it to bend *away* from the low dielectric region.[49] Although the neutralised ligand does generate electric field gradients, its enhanced negative surface charge should also accentuate the phosphate repulsions and induce the DNA to bend away from the ligand, or at least coil less tightly about it. The fact that this is not observed, and indeed the coiling may be enhanced for the neutralised ligand, implies that it is the other forces—the van der Waals forces—that drive the coiling.

Ligand charge does have some affect on the DNA response, but its principal manifestation is different in the two force fields. The contrast is instructive. In CHARMM27, the $+4\,e$ charge attracts the negatively charged DNA backbone, causing the major groove to contract onto the cylinder. With CHARMM22 its effect is to strain the hydrogen bonding between base pairs in the duplex. The induction of such strain is reasonable given that most force fields describe hydrogen bonding primarily through electrostatic interactions. Orbital interactions, such as charge transfer and π-electron polarisation, do contribute strongly to the hydrogen bonding between base pairs,[50,51] but these are also likely to be strongly perturbed by the proximity of a tetracationic ligand. In reality, both the strain to the base pair hydrogen bonds and the contraction of the DNA backbone onto the ligand are likely to be present, and getting the balance right will be an important validation of the force field. In this context, the CHARMM22 result that duplex base-pairing will be disrupted by C^{4+} provides an interesting prediction that awaits the measurement of definitive X-ray or NMR structures.

Sensitivity to cylinder charge: If, as indicated above, the disruption of base pairs in a DNA duplex with CHARMM22 is due to a competition between inter-base pair hydrogen bonding and ligand base electrostatics, then it is of interest to identify how strong the ligand electrostatics need to be in order to effect the base pair mismatches. Analogous simulations to those reported in the last section were performed with CHARMM22 and with a variant of the cylinder in which the overall charge was set to $+3\,e$ (C^{3+}), again achieved by a constant shift of the atomic charges for all C and non-polar H atoms. Within 0.25 ns this system also showed disruption of the base pairing within the double helix (see Figure 15). Similar calculations with a $+2\,e$ variant of the cylinder (not shown) gave results that were intermediate between C^{3+} and C^0, with frequent bifurcated pairings in which an A was found to bridge two T bases in the opposite strand, but these did not lead to irreversible breakdown in the base pairing on a nanosecond timescale. We conclude that some disruption of the duplex base pairs is a robust feature of the CHARMM22 model, and is not especially sensitive to the parameterisation of the cylinder charge distribution.

Conclusion

Our molecular dynamics simulations of the supramolecular cylinder, **1**, in explicit solvent show major groove binding of the cylinder on the DNA and coiling of the DNA in re-

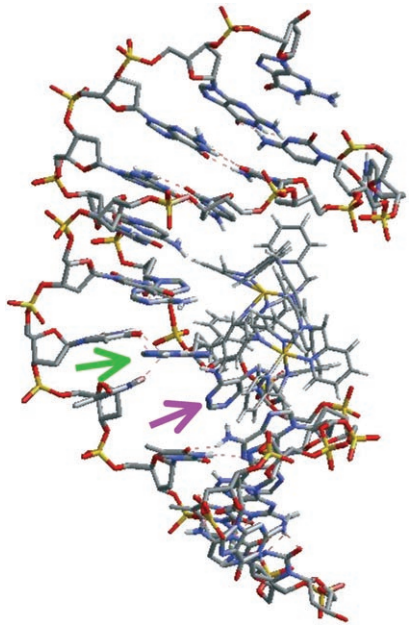


Figure 15. Configuration of DNA 0.25 ns after continuing from the final configuration of the 2 ns DNA/C^0 simulation with a $+3\,e$ cylinder (C^{3+}). The magenta arrow points to an unpaired T, while the green arrow points to two A bases paired to a single T.

sponse to this binding. This is consistent with all available experimental observations. Moreover, the simulations have, as envisaged, yielded molecular level information about the coiling effect. Other highly charged cations, such as spermine and cobalt(III)–hexamine,[17,22,52,53] bend DNA by a few degrees per ligand and cause it to condense into poly-molecular aggregates (often of well-defined shape). Our simulations indicate that the DNA bend induced by **1** is an order of magnitude greater than this: the simulations give 20–40°, depending on the force-field model, which compares favourably with the experimental measurements of 40–60° per ligand. Somewhat surprisingly, the tetracationic charge of the supramolecular cylinder was found *not* to be essential for inducing the coiling, as a neutral analogue was also found to bend the DNA. Instead the molecular shape appears to be the key factor and must give rise to a coil with a pitch which ensures an intramolecular effect observed experimentally with this ligand, rather than the more commonly observed intermolecular aggregation. This highlights the significance of our molecular design which creates, through metallo-supramolecular assembly, a cylindrical agent of very similar shape and dimensions to the various cylindrical motifs employed by proteins for DNA recognition and structure control.

This is *not* to say electrostatic interactions were unimportant. Our simulations identified two competing influences arising from a cationic ligand in the major groove that are likely to be general effects. On the one hand, the large cationic charge attracted the phosphate backbone, causing the major groove to contract around the ligand. On the other hand, such a large positive charge density close to the base pairs places strain upon the inter-base hydrogen bonding that holds the DNA duplex together. The balance of these two effects did depend on the choice of DNA force field, but the latter effect was sufficiently large that C^{4+} was observed to generate base-pair mismatches in some systems. When these mismatches were observed, they occurred in the A–T tract and adjacent to, rather than at, the site at which the cylinder bound. Any weakening of the base pair hydrogen bonding at or near this site could prime the DNA for transcription or replication processes. This is an unexpected molecular-level feature associated with the experimentally observed cylinder binding and DNA conformational change. Intriguingly, we note that many proteins also bend DNA at locations adjacent to the site of transcription. For example, the CAP (catabolic gene activating protein) binds adjacent to the initial site at which RNA polymerase acts and causes a 50-fold increase in the rate of transcription initiation.[54] This activation is usually ascribed to the provision of an interaction surface for the polymerase, thereby enhancing binding to the promoter site. Our C^{4+}/DNA simulations suggest there may be an alternative mechanism: that the bending induced by CAP binding can help to open the DNA at an adjacent site, and could thus contribute to the increased rate at which transcription is initiated in the presence of promoters. This opens up new potential applications for these DNA-coiling metallo-supramolecular cylinders.

Thus we conclude that the tetracationic supramolecular cylinder, **1**, does have two substantive effects on DNA. In the first place it can induce coiling of the DNA, with potentially very large curvature being introduced in DNA dodecamers. However, at the same time it weakens the hydrogen bonding between base pairs near the binding site and so can tend to introduce defects into the alignment of the base-pair sequences in the two strands of the double helix. In our simulations this occurred consistently within the A–T rich region of the DNA and resulted in misaligned A–T cross-links between the two strands. It is not clear whether this represents a selectivity for A–T over G–C and simulations with other DNA sequences are in progress to resolve this issue.[1,2]

Acknowledgements

We thank the EPSRC (S.K.; postgraduate studentship) and the European Union (MARCY RTN; HPRN-CT-2002-00175) for support. M.J.H. is the Royal Society of Chemistry Sir Edward Frankland Fellow 2004-5.

[1] See, for example, J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th ed., Freeman, New York, **2002**.
[2] J. P. Baak, F. R. Path, M. A. Hermsen, G. Meijer, J. Schmidt, E. A. Janssen, *Eur. J. Cancer* **2003**, *39*, 1199–1215.
[3] C. F. Calkhoven, C. Muller, A. Leutz, *Trends Mol. Med.* **2002**, *8*, 577–583.
[4] L. Rogge, *Arthritis Res. Ther.* **2003**, *5*, 47–50.
[5] B. S. Parekh, G. W. Hatfield, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1173–1177.

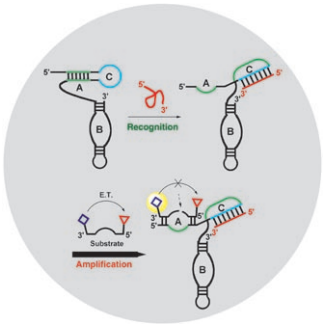
[6] R. R. Sinden, *DNA Structure and Function*, Academic Press, London, **1994**.
[7] R. E. Dickerson, *Nucleic Acids Res.* **1998**, *26*, 1906–1926.
[8] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland, New York, 2nd ed., **1999**.
[9] B. C. Baguley, *Mol. Cell. Biochem.* **1982**, *43*, 167–181.
[10] W. A. Denny, *Anti-Cancer Drug Des.* **1989**, *4*, 241–263.
[11] S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA, **1994**.
[12] D. Z. Coggan, I. S. Haworth, P. J. Bates, A. Robinson, A. Rodger, *Inorg. Chem.* **1999**, *38*, 4486–4497.
[13] B. Onfelt, P. Lincoln, B. Nordén, *J. Am. Chem. Soc.* **1999**, *121*, 10846–10847.
[14] K. E. Erkkila, D. T. Odom, J. K. Barton, *Chem. Rev.* **1999**, *99*, 2777–2795.
[15] M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 879–884.
[16] I. Meistermann, A. Rodger, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, S. Khalid, P. M. Rodger, J. Peberdy, C. J. Isaac, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074.
[17] A. Rodger, K. J. Sanders, M. J. Hannon, I. Meistermann, A. Parkinson, D. S. Vidler, I. S. Haworth, *Chirality* **2000**, *12*, 221–236.
[18] G. S. Manning, K. K. Ebralidse A. D. Mirzabekov, A. Rich, *J. Biomol. Struct. Dyn.* **1989**, *6*, 877–889.
[19] E. Moldrheim, M. J. Hannon, I. Meistermann, A. Rodger, E. Sletten, *J. Biol. Inorg. Chem.* **2002**, *7*, 781–790.
[20] I. S. Haworth, A. Elcock, A. Rodger, W. G. Richards, *J. Biomol. Struct. Dyn.* **1991**, *9*, 553–569.
[21] B. P. Hudson, J. K. Barton, *J. Am. Chem. Soc.* **1998**, *120*, 6877–6888.
[22] T. E. Cheatham, P. A. Kollman, *Structure* **1997**, *5*, 1297–1311.
[23] Y. Y. Fang, B. D. Ray, C. A. Claussen, K. B. Lipkowitz, E. C. Long, *J. Am. Chem. Soc.* **2004**, *126*, 5403–5412.
[24] B. Wellenzohn, W. Flader, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *Biopolymers* **2001**, *61*, 276–286.
[25] J. Dolenc, C. Oostenbrink, J. Koller, W. F. van Gunsteren, *Nucleic Acids Res.* **2005**, *33*, 725–733.
[26] A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnik, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
[27] N. Foloppe, A. D. MacKerell, Jr., *J. Comput. Chem.* **2000**, *21* 86–104.
[28] S. Y. Reddy, F. Leclerc, M. Karplus, *Biophys. J.* **2003**, *84*, 1421–1449.
[29] W. Smith, T. R. Forester, *J. Mol. Graphics* **1996**, *14*,136.
[30] a) M. A. San Miguel, R. Marrington, P. M. Rodger, A. Rodger, C. Robinson, *Eur. J. Biochem.* **2003**, *270*, 3345–3352; b) M. A. San Miguel, P. M. Rodger, *Phys. Chem. Chem. Phys.* **2003**, *5*, 575–581; c) R. Lukac, A. J. Clark, M. A. San Miguel, A. Rodger, P. M. Rodger, *J. Mol. Liq.* **2002**, *101*, 261–272; d) T. Astley, G. G. Birch, M. G. B. Drew, P. M. Rodger, *J. Phys. Chem. A* **1999**, *103*, 5080–5090.
[31] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
[32] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comp. Physiol.* **1977**, *23*, 327.
[33] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford (UK), **1987**.
[34] M. Feig, B. M. Pettitt, *Biophys. J.* **1998**, *75*, 134–149.
[35] N. Korolev, A. P. Lyubartsev, A. Laaksonen, L. Nordenskiöld, *Biophys. J.* **2002**, *82*, 2860–2875.
[36] S. Khalid, PhD Thesis, University of Warwick (UK), **2004**.
[37] B. Wellenzohn, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *J. Am. Chem. Soc.* **2000**, *122*, 3927–3931.
[38] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 655–667.
[39] G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 669–699.
[40] R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. M. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C. S. Tung, A. H. J. Wang, V. B. Zhurkin, *EMBO. J.* **1989**, *8*, 1–4.
[41] W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinmann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, H. M. Berman, *J. Mol. Biol.* **2001**, *313*, 229–237.
[42] X. J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
[43] D. L. Beveridge, S. B. Dixit, G. Barreiro, K. M. Thayer, *Biopolymers* **2004**, *73*, 380–403; D. McDonald, P. Lu, *Curr. Opin. Struct. Biol.* **2002**, *12*, 337–343.
[44] R. E. Dickerson, T. K. Chiu, *Biopolymers* **1997**, *44*, 361–403.
[45] R. R. Sinden, *DNA Structure and Function*, Academic Press, New York, **1994**.
[46] A. Barbic, D. P. Zimmer, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2369–2373.
[47] J. Hizver, H. Rozenberg, F. Frowlow, D. Rabinovich, Z. Shakked, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8490–8495.
[48] M. A. Hassan, C. R. Calladine, *J. Mol. Biol.* **1998**, *282*, 331–343.
[49] A. H. Elcock, A. McCammon, *J. Am. Chem. Soc.* **1996**, *118*, 3787–3788.
[50] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **2002**, *114*, 2194–2197; *Angew. Chem. Int. Ed.* **2002**, *41*, 2092–2095.
[51] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111*, 3120–3122; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942–2945.
[52] I. S. Haworth, A. Rodger, W. G. Richards, *Proc. R. Soc. London Ser. B* **1991**, *244*, 107–116.
[53] K. Bryson, R. J. Greenall, *J. Biomol. Struct. Dyn.* **2000**, *18*, 393–412.
[54] See reference [1], p. 873.

Received: September 22, 2005
Published online: February 23, 2006

Cover Picture

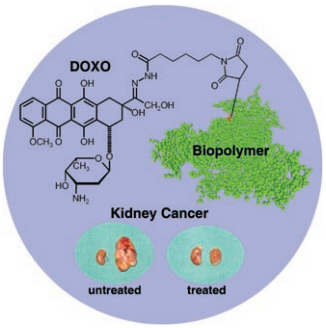
Aneta Oleksi, Alexandre G. Blanco, Roeland Boer, Isabel Usón, Joan Aymamí, Alison Rodger, Michael J. Hannon,* and Miquel Coll*

Metallosupramolecular helicates with trigonal-antiprismatic geometry thread through and fit perfectly into the central hydrophobic cavity of a three-way DNA junction. Six hydrophobic rings of the drug stack on six DNA bases, with three more sandwiched in the minor grooves. This work reveals and enables a new approach to targeting DNA and specifically DNA junctions. For more details, see the Communication by M. Coll, M. J. Hannon, and co-workers on page 1227 ff.



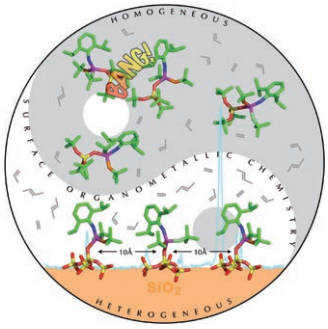
Sensors

The molecular recognition between an external signal and a receptor is transduced allosterically to catalytically amplified chemical information. E. V. Anslyn and L. Zhu outline this approach in their Minireview on page 1190 ff.



Medicinal Chemistry

Polymeric carrier systems enable the optimization and transport of low-molecular-weight drugs, proteins, and oligonucleotides. Current developments in polymer therapeutics are given by R. Haag and F. Kratz in their Review on page 1198 ff.



Metathesis Catalysts

In their Communication on page 1216 ff., C. Copéret and co-workers describe the preparation of a silica-supported Mo–carbene catalyst for olefin metathesis and how its performance compares relative to its molecular equivalent.

DOI: 10.1002/chem.200501168

Simulations of DNA Coiling around a Synthetic Supramolecular Cylinder That Binds in the DNA Major Groove

Syma Khalid,[a, c] Michael J. Hannon,[a, b] Alison Rodger,[a] and P. Mark Rodger*[a]

Abstract: In this work we present the results of a molecular simulation study of the interaction between a tetracationic bis iron(II) supramolecular cylinder, $[Fe_2(C_{25}H_{20}N_4)_3]^{4+}$, and DNA. This supramolecular cylinder has been shown to bind in the major groove of DNA and to induce dramatic coiling of the DNA. The simulations have been designed to elucidate the interactions that lead the cylinder to target the major groove and that drive the subsequent DNA conformational changes. Three sets of multi-nanosecond simulations have been performed: one of the uncomplexed d(CCCCCTTTTTCC)·d(GGAAAAAGGGGG) dodecamer; one of this DNA complexed with the cylinder molecule; and one of this DNA complexed with a neutralised version of the cylinder. Coiling of the DNA was observed in the DNA–cylinder simulations, giving insight into the molecular level nature of the supramolecular coiling observed experimentally. The cylinder charge was found not to be essential for the DNA coiling, which implies that the DNA response is moderated by the short range interactions that define the molecular shape. Cylinder charge did, however, affect the integrity of the DNA duplex, to the extent that, under some circumstances, the tetracationic cylinder induced defects in the DNA base pairing at locations adjacent to the cylinder binding site.

Keywords: bio-supramolecular chemistry · DNA recognition · helical structures · molecular dynamics · noncovalent interactions

Introduction

Control of gene expression is currently one of the key areas of interest in molecular medicine.[1–4] Gene expression involves the transfer of information encoded within the gene to produce a biologically active protein. However, not all genes are expressed in every cell all the time; bio-regulation of gene expression is effected by proteins that activate or repress transcription by binding to short, specific DNA sequences.[1,5] The ability to turn genes on or off artificially by the action of synthetic analogues of DNA-binding proteins is an important goal that would open up new possibilities for disease control and prevention as well as cure.

Proteins that bind DNA frequently achieve sequence-specific code recognition by binding non-covalently in the major groove of DNA.[6–8] In biological systems, the major groove is the preferred binding site for sequence recognition as this groove shows the greater variation in size and shape with base sequence and also the greater number and variation in pattern of hydrogen-bond donor and acceptor units; it is therefore the ideal target for synthetic molecules designed to recognise and bind to specific DNA sequences. However, relatively little progress had been made in producing synthetic major groove binders because of the size entailed. Most of the compounds synthesised are either minor groove binders,[9] intercalators,[10,11] or metal complexes that span only two to three base pairs and whose interaction with DNA is too limited for sequence selectivity.[12–14] Design of synthetic agents that target the major groove in a non-covalent, sequence-selective fashion remains a challenge of considerable importance in molecular medicine and biology.

[a] Dr. S. Khalid, Prof. M. J. Hannon, Prof. A. Rodger, Prof. P. M. Rodger
Department of Chemistry, University of Warwick
Coventry CV4 7AL (UK)
E-mail: p.m.rodger@warwick.ac.uk

[b] Prof. M. J. Hannon
Current address:
School of Chemistry, University of Birmingham
Birmingham B15 2TT (UK)

[c] Dr. S. Khalid
Current address:
Department of Biochemistry, University of Oxford
South Parks Rd, Oxford, OX1 3QU (UK)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

Recently, a major step has been taken toward achieving the goal of designing synthetic agents to target the major groove: we have developed a novel compound that binds strongly in the major groove of DNA and is large enough to span five base pairs.[15,16] The compound is a metallo-supramolecular tetracationic cylinder (**1**) with a triple helical framework. This cylinder is approximately 19 Å in length and 11 Å in diameter. It is too big to bind in the minor groove of DNA, but has just the right shape and size to lie along the major groove.[15]

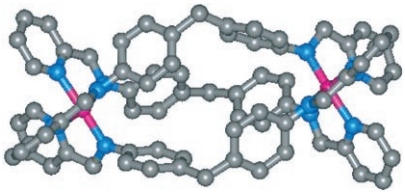


Figure 1. Tetracationic cylinder $[Fe_2L_3]^{4+}$, where L is *N*,*N*′-bis(pyridin-2-ylmethylene)-4,4′-diaminodiphenylmethane: Fe^{II} is pink, N is blue, C is grey; H atoms have been omitted for clarity.

Experimental techniques have provided information regarding the binding strength and preferred binding sites of **1** on DNA.[16] Spectroscopic studies indicate that the cylinder binds very strongly to DNA with a binding constant well in excess of $10^7\,\text{M}^{-1}$ in 20 mM salt. Flow linear dichroism (LD) reveals coiling or bending of the DNA on addition of the cylinder, and tapping mode atomic force microscopy (AFM)[16] reveals this to be dramatic intramolecular coiling of DNA that is unprecedented with a synthetic agent: **1** was found to induce DNA to bend by 40–60° per ligand,[15] with the smaller angles (40°) occurring at the maximum loading of one cylinder per DNA turn. In contrast cobalt ammines—which are known to cause DNA to bend and ultimately to condense into multi-molecular aggregates—generate just 2–5° bend per ligand.[17] The effect of cobalt ammines can largely be explained by the Manning charge condensation theory,[18] but the magnitude of the effect with **1** suggests the bending is of a very different character. The *M*-helical enantiomer of the cylinder coils DNA more aggressively than the *P* enantiomer.[16] NMR data have confirmed a major groove-binding mode for the *M* cylinder.[19] From all the experimental data it is clear that the cylinder binds in the major groove and is able to induce dramatic conformational changes in the DNA. However, gaining molecular level information about the effects and interactions which lead to this remarkable supramolecular event is challenging. Molecular dynamics (MD) simulations can provide information at the molecular level that is complementary to experiment and therefore are an ideal way to get a better understanding of this system.

There is comparatively little modelling literature on the interaction between transition-metal complexes and DNA. Some modelling has been performed on DNA with ruthenium(II)–tris(1,10-phenanthroline)[20] in vacuum, and the results compared with experimental data.[12] The Λ-ruthenium complex showed a preference for the major groove of DNA but little or no bending of the DNA at low loading and whereas the Δ enantiomer preferred the minor groove at low loadings where it did bend the DNA—but only by a few degrees—and bound to both grooves at higher loadings with no additional bending. Several other papers have used modelling studies to interpret experimental data on DNA/transition-metal complexes,[17,21] but since their emphasis was on understanding the experimental data, these calculations again omitted the solvent. Rigorous molecular dynamics simulations with explicit solvent have been performed for cobalt(III)–hexammine[22] and a nickel(II)–metalloprotein[23] in the presence of DNA. In both cases, the simulations produced good agreement with a range of experimental data; in neither case was there significant bending or coiling of DNA.

In this paper we present results of a molecular dynamics (MD) study of the effect of the supramolecular cylinder, **1**, on DNA. Docking calculations, both manual and using high-temperature MD with constrained or rigid molecules, have been used to identify favourable initial binding sites for the cylinder on a DNA dodecamer, and then multi-nanosecond MD simulations performed with explicit solvent to monitor the DNA response. Previously, MD simulations[19] were performed on this system to interpret the NMR spectra, but these used NOE data as additional force-field restraints. The simulations described in the present paper were designed to probe any changes in DNA structure upon cylinder binding and are not restrained in any way. As a consequence, these new simulations enable a direct analysis of the DNA response—both its nature and the forces that give rise to it.

Computational Methods

System: The DNA dodecamer sequence (CCCCCTTTTTCC)·d(GGAAAAAGGGGG) was chosen for this study. Preliminary simulations were performed with a decamer, d(CCCCCTTTTT)·d(AAAAAGGGGG), which was chosen since it might offer potential for identifying preference of A–T versus G–C tracts within a single series of simulations; this sequence was subsequently extended by adding two C–G pairs to the A–T end so as to minimise any tendency for the DNA ends to fray in the absence of the cylinder; note that no restraints were introduced into the force field to hinder end-fraying. The numbering used in this paper to identify specific bases and base pairs is defined below.

5′	C1	C2	C3	C4	C5	T6	T7	T8	T9	T10	C11	C12	3′
3′	G24	G23	G22	G21	G20	A19	A18	A17	A16	A15	G14	G13	5′
base pair no.	1	2	3	4	5	6	7	8	9	10	11	12	

The *M* enantiomer of the cylinder, **1** (denoted C^{4+}), was modelled using the CHARMM22 all-atom force field, but with the FeN_6 sub-unit treated as a rigid body with geometry taken from the crystal structure. For comparison, a simulation was also performed in which the *M* cylinder was

made electrically neutral, denoted C^0. Although the C^0 system has no experimental counterpart, a comparison of the C^{4+} and C^0 systems enables an exploration of the relative importance of molecular shape (as defined by the van der Waals interactions) and electrostatic interactions in inducing the DNA response. Analogous comparisons with netropsin[24,25] have proved useful in understanding the influence of charge in minor-groove binding ligands. Many methods could be used to achieve this neutrality, including setting all atomic charges to zero, neutralising the Fe atoms, or adding a neutralising negative charge to the organic part of the cylinder. In this work we have adopted the last option: $-0.03\,e$ was added to the charge on each C and H atom in the cylinder, resulting in a counterbalancing $-4\,e$ charge spread over the surface of the cylinder. While still giving rise to electric field gradients, the negative surface charge should considerably reduce the *electrostatic* attraction of the cylinder to the DNA phosphate backbone, and thus makes the comparison with C^{4+} instructive.

Technical details: Compounds were modelled using the CHARMM force-field series. This contains two main variants for modelling DNA: the older CHARMM22 force field,[26] and a more refined CHARMM27.[27] There have been a number of comparisons of how well these two force fields reproduce the behaviour of DNA,[28] and the general consensus is that the CHARMM27 force field is much better. In particular, it correctly predicts B-DNA to be the stable conformation in low ionic strength solvents at normal temperatures, whereas the CHARMM22 force field tends to cause the DNA to adopt an A-like form. However, it should be noted that most of these comparisons have treated just the duplex in aqueous solution: no comparisons have been reported for DNA interacting with a major-groove-binding ligand, particularly when that ligand carries a substantial positive charge. The experimental data discussed above indicates that C^{4+} perturbs the DNA strongly, bending it well away from the canonical B-form. This is outside the parameter-space that has so far been used to derive and validate the DNA force fields, and so it is of interest to compare the performance of these two force fields in modelling DNA/C^{4+} complexes.

All simulations were performed with DL_POLY.[29] The conversion from CHARMM to DL_POLY force-field formats was achieved by using a purpose-built program that interprets the CHARMM prm and crd files. As in previous work[30] checks on the force field were performed on numerous configurations to ensure the energies and forces calculated with DL_POLY and CHARMM agreed *exactly*. Water was modelled with the TIP3P[31] potential, and kept rigid using the SHAKE algorithm,[32] implemented in DL_POLY with a tolerance of 0.0001. All hydrogen atoms were assigned a mass of 2 u. This gave good energy conservation in constant energy simulations with a 2 fs time step for the CHARMM22 force field. An Ewald sum[33] was used to evaluate electrostatic interactions, with $\boldsymbol{k}_{max} = (5,5,6)$ and $\alpha = 0.12604\ \text{Å}^{-1}$. Simulations were performed with orthorhombic periodic boundaries at constant volume and temperature (NVT) using the Hoover thermostat[33] with a time constant of 0.1 ps and a time step of 2 fs; some 0.25 ns segments of the trajectories were repeated with a 1 fs time step and showed no significant differences. Some difficulties were encountered initiating CHARMM27 simulations with the 2 fs time step, and so this was reduced to 1 fs for all the CHARMM27 simulations reported herein. In all other respects, the CHARMM22 and CHARMM27 simulations followed the same protocols and methodology.

Protocol: The multi-nanosecond simulations were performed using a four-step protocol designed to identify good binding sites while still ensuring a disperse background ion distribution:

1) High temperature MD simulations were performed on the cylinder and ions moving around an immobilised B-DNA dodecamer, with the relative dielectric constant $\varepsilon_r = 80$;
2) low potential energy configurations were identified and optimised with respect to just the cylinder position; ions were kept fixed at their positions from the high temperature snapshot and the DNA also remained fixed;
3) water was added, the relative dielectric constant reset to 1, and the system allowed to relax while restraining the DNA with a series of progressively weaker harmonic tethering potentials;
4) full, unrestrained MD simulations were performed on the aqueous system for 2–5 ns.

This protocol is described more fully hereafter.

Docking (steps 1 and 2): A dodecamer of B-DNA and one cylinder molecule (*M* enantiomer) were embedded in a neutralising atmosphere of Na^+ ions. A further 58 Na^+ and 58 Cl^- ions were added in a $45\ \text{Å} \times 45\ \text{Å} \times 60\ \text{Å}$ box, with DNA aligned along the z (long) axis; this gave [NaCl] = 0.8 M, which is similar to the concentration used in some other MD studies.[34,35] The DNA was immobilized and an NVT MD simulation performed for 1 ns at 900 K; this temperature was found to be high enough to ensure that the ligand sampled the whole surface of the DNA efficiently, but without any significant deformation in its own shape. The temperature was also high enough to ensure that the background atmosphere of sodium and chloride ions behaved like a homogeneous gas, with no aggregation of the ions onto the DNA or cylinder. The conformations with the lowest configurational energy were extracted and energy-minimised with respect to the position of just the ligand. No change was allowed in the position of the DNA atoms, Na^+ ions or Cl^- ions during this optimisation of the cylinder within its DNA/ionic gas environment. The resulting configuration was then used to start fully solvated MD simulations as described below. Variants on this procedure were tried in which the DNA and/or Na^+ and Cl^- ions were harmonically restrained rather than frozen during the cylinder minimisation, but these produced essentially the same results when carried forward into fully solvated MD simulations. Manual docking calculations were also performed, but did not give more favourable docking sites than those identified using this high temperature MD docking protocol.

A number of alternative low energy conformations was examined and all found to involve major groove binding. Several of these were then carried forward into full solvated MD simulations following the protocol given below.[36] All such simulations exhibited a very similar DNA response to that reported herein—particularly with respect to the extent to which they bent the DNA, and the resultant stability of the DNA—and confirmed that the results of our simulations were not sensitive to variations in the initial binding site. We note that these repeat simulations also proved our results to be robust with respect to the initial background ion distribution, since each configuration also involved a very different, essentially random, initial arrangement of the sodium and chloride ions.

Molecular dynamics simulations (steps 3 and 4): Water molecules, taken from an equilibrated liquid water simulation, were added to the optimised configurations (DNA, cylinder, Na^+ and Cl^- positions) identified in the docking calculations. Any water molecule that overlapped with DNA, cylinder, Na^+ or Cl^- was removed, which left a total of 3720 water molecules in the final DNA/ligand system, or 3758 for the uncomplexed DNA. Equilibration followed a similar protocol to that used by other research groups.[23,28,37] The DNA atoms were tethered to their original positions with a harmonic force constant of $100\ \text{kcal mol}^{-1}\ \text{Å}^{-2}$ and an NVT MD simulation performed for 10 ps at 310 K. A further five 10 ps simulations were then performed successively in which the tethering force constant was 50, 25, 10, 5 and $1\ \text{kcal mol}^{-1}\ \text{Å}^{-2}$, respectively. The tethering potentials were then removed completely and multi-nanosecond simulations performed, saving configurations every 1 ps for later analysis.

Up to four repeat simulations of 2–5 ns duration, using slightly different starting configurations, were performed for each ligand/DNA system to validate the reproducibility and timescale of the DNA response. In every case, the response of the DNA was found to be rapid and repeatable, with most changes occurring in the first 0.5 ns, and no noticeable changes occurring between 2 and 5 ns.[36] In this paper we therefore concentrate on the behaviour during the first 2 ns after binding.

Analysis methods: A number of different methods have been used in this work to monitor DNA structure. The time dependence of conformational parameters has been monitored using the Curves algorithm[38] implemented in MDToolchest.[39] This uses seven torsion angles (α–ζ, ϕ, χ) to describe the DNA backbone, while the intra-base-pair geometry is described by six *helicoidal* parameters: three displacements (shear,

stretch and stagger) and three angles (buckle, propeller and opening). This parameter set is now well established as a means of describing DNA conformation;[40,41] a full definition is supplied with the Supporting Information, though it is useful to note here that of the backbone angles, α and ζ refer to torsions about P–O bonds, δ and ϕ refer to the ribose ring, while χ is for the bond that links the backbone to the base.

In understanding the DNA conformation and flexibility, it is also useful to examine some of the parameters that give a more correlated view of the overall behaviour. In particular, it is useful to study the relationship between neighbouring base pairs. In this work these have been monitored using a distance (slide) and three angles (tilt, roll and twist; see Figure 9) as implemented in 3DNA.[42] In this case, the analysis has been applied to the average structure determined from a continuous 50 ps portion of the MD trajectory; this length of trajectory was found to be long enough to smooth out the instantaneous fluctuations in shape, but was still short with respect to the systematic relaxation induced by the presence of the cylinder, **1**. The 3DNA analysis of average structures has also been used to generate normal vector plots,[7,43] which are useful for identifying linear and bent regions within the DNA, and hence to estimate the degree to which the DNA is bent by the cylinder.

Results and Discussion

Response of the DNA to a bound cylinder

As will be discussed below, the simulations with the two DNA force fields (CHARMM22 and CHARMM27) gave broadly similar responses for the DNA/ligand complex, in that the binding site was consistent with the two models and both indicated the DNA to be bent by the cylinder. Some quantitative differences were observed between the two force fields, with CHARMM27 tending to give the more minimal response to the cylinder. Two examples (discussed in more detail below) are worthy note at this stage. The overall bend of the DNA was about 40° with CHARMM22 compared with 20° for CHARMM27, and the CHARMM22 DNA did respond differently to C^{4+} and C^{0} whereas the CHARMM27 model was remarkably insensitive to this large change in ligand charge. In the following we focus first on the results for the CHARMM22 model: since the response is greater with this model it is likely to be easier to identify the underlying physics of the DNA/ligand interaction. This is then followed by a more detailed comparison of the force fields to place confidence limits on the interpretation of the simulations.

Charmm22: As described in the protocol, favourable binding sites for the cylinder on the DNA were identified and then explicit solvent molecules added. In all cases, binding was found to occur preferentially in the region of the A–T tract (base pairs 6–10). Using the position from the docking calculation as a starting point, water was added, the system equilibrated while the DNA conformation was restrained to the B-form, and then simulations allowed to proceed without restraints. We re-iterate that up to four such simulations were performed for each DNA/cylinder system.

Once the restraints were removed, the DNA was observed to respond rapidly to the presence of the cylinder. Major conformational changes were observed in the presence of both C^{4+} and C^{0}. In every trajectory these occurred within the first 0.5 ns, with the precise timing depending on the starting configuration. In all cases, no significant conformational changes were observed in the subsequent 4–5 ns, suggesting that our simulations were probing the final DNA state, or at least identifying a long-lived intermediate in the response process.

It is instructive to begin by comparing before and after images of the DNA conformation in the three different systems: uncomplexed, bound to C^{4+}, and bound to C^{0}. Although such snapshots provide only limited information about the range of DNA conformations, for the present study they prove to give a useful and visual overview of the more quantitative data presented below. Images of the initial (directly after docking) and relaxed (after 2 ns) configurations for DNA and cylinder are given in Figure 2. Initially, the DNA adopts a nearly linear B-form. The initial binding mode of the cylinder is in the major groove, lying symmetrically between the two strands of the DNA, and spanning base pairs 6–11 (T6–A19 to C11–G14). This is consistent with the published structure obtained from NMR (NOE) data.[15,16,19] This binding geometry was retained throughout the equilibration phases, and was not disrupted until the tethering potentials that restrained the DNA conformation were removed.

After 2 ns of the unrestrained simulations, the *uncomplexed* DNA retained its near linear B-form, albeit with some fraying at the first base pair. In contrast, the final configurations of the DNA/cylinder simulations suggest substantial curvature of the DNA for both charged and uncharged cylinders. As will be discussed below (see Figure 7), these snapshots do depict a bend of about 40° in the helix axis in the presence of the cylinders.

At the end of the simulations the charged cylinder, C^{4+}, remained in the major groove, but some distortion of the groove shape was evident. There was close association between the cylinder and base-pairs 5–9 (closest interatomic distances between the cylinder and each of these nucleic acids are 2.4–3.5 Å); thus the cylinder remained within the A-T tract. There was no evidence of hydrogen bonding between the cylinder and DNA, but this is not surprising since the cylinder has no conventional hydrogen-bond donor groups, and its potential acceptors are obscured by nonpolar carbons. It is particularly interesting to note that some breakdown of the Watson–Crick structure is apparent in the last three or four base pairs of the DNA (Figure 2, middle, bottom), that is, adjacent to the cylinder binding site. As is shown below, this is due to a mispairing between T9 and A15, which then leaves T10 and A16 unpaired.

Importantly, the simulations reproduce both the major groove binding characterized by NMR and the DNA coiling effects as observed in LD and AFM experiments. The implication of the simulations is that, at least for some sequences, this bending might be associated with alterations to the DNA base pairing. Such effects have also been observed in some DNA–protein complexes,[44,45] and can, indeed, be an important feature of the way that proteins process DNA.

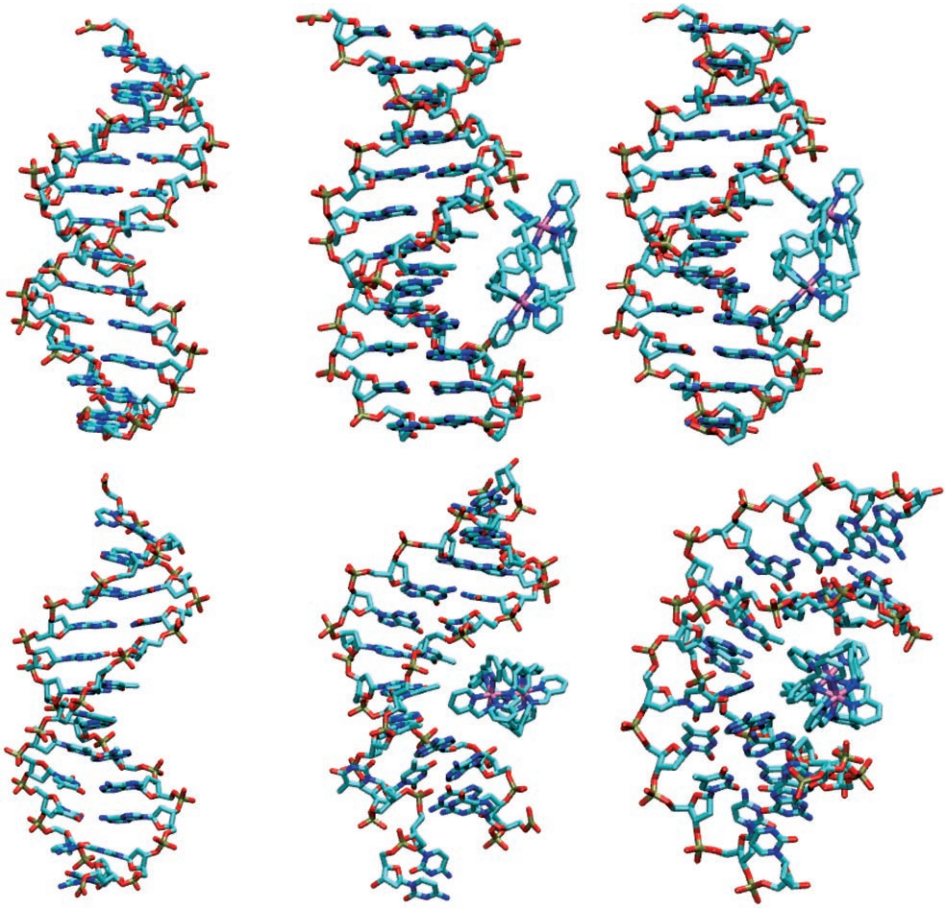


Figure 2. Conformations of DNA taken from the beginning (top) and after 2 ns (bottom) of MD simulations using the CHARMM22 force field: uncomplexed DNA (left), DNA/C^{4+} (middle) and DNA/C^{0} (right). Conformations after 5 ns were very similar to those at 2 ns.

Moreover, DNA bending is commonly associated with A–T tracts.[46,47]

The behaviour of the neutralised cylinder, C^{0}, provides a fascinating contrast to that of C^{4+}. Its final configuration shows C^{0} to lie symmetrically within the major groove, spanning base pairs 5–10, but with the DNA having bent substantially to create a much deeper pocket for the cylinder than was found with C^{4+}. It might originally be thought that the coiling of the DNA was, in large part, due to the attraction between the large positive charge on C^{4+} and the negatively charged phosphates in the DNA backbone. However, it is clear from Figure 2 (and the quantitative analysis presented below) that the neutral cylinder, C^{0}, also causes strong coiling in the DNA, but does so without inducing any of the disruption of the Watson–Crick base pairs that was evident with C^{4+}. This contrast suggests that the shorter ranged van der Waals forces play a significant, perhaps even dominant, role in coiling the DNA backbone while the large Coulombic forces generated by the exposed cationic charge of C^{4+} may generate a localised stress that is strong enough to disrupt inter-strand hydrogen bonding in the DNA duplex.

The helicoidal parameters developed by Lavery and Sklenar[38] are a good source of quantitative data on how these ligands affect the integrity of the base pairs within the DNA double helix. These parameters describe how two nucleic acid bases move and rotate relative to each other and are defined such that zero displacement/rotation corresponds to the ideal base pair geometry for B-DNA. Helicoidal parameters have been calculated from all three simulations and are presented in Figure 3. For the uncomplexed DNA, nearly all the base pairs can be seen to be stable, with typically small fluctuations about values of zero (the ideal base-pair geometry). The only exception is the first pair, C1–G24, which, from about 1 ns, showed large amplitude motions characteristic of fraying at the end of the double helix; end-fraying is a not uncommon event for DNA, both in simulations and in vivo. Some large-amplitude fluctuations were also seen for T8–A17 and T9–A16 pair at about 1 ns, but these were transient and rapidly returned to stable values; such behaviour is indicative of the flexibility inherent within the DNA duplex rather than of irreversible conformational changes.

Given the amount of curvature evident in Figure 2 (right), the C^{0} ligand gives rise to helicoidal parameters that are remarkably similar to those of the uncomplexed DNA. Indeed, the principal difference between these two sets of data is that the transient large amplitude oscillations seen in the uncomplexed DNA are completely absent for DNA in the presence of C^{0}. As with the uncomplexed DNA, fraying is again found only at the first base pair. Thus, the presence of C^{0} actually appears to enhance the stability of the Watson–Crick base pairs, despite the overall curvature of the DNA evident in Figure 2 (right).

For C^{4+} bound to DNA, the helicoidal parameters clearly reveal a deformation of the base-pair stack. There are two distinct zones in the double helix for this system. The first seven base pairs show stable behaviour as seen in the uncomplexed DNA, with just small fluctuations around zero. Similar behaviour is seen for T8–A17, albeit with larger amplitude vibrations for the shear (SHR), buckle (BKL) and open (OPN) parameters. However, the plots for the last four base pairs show extremely large variations that are

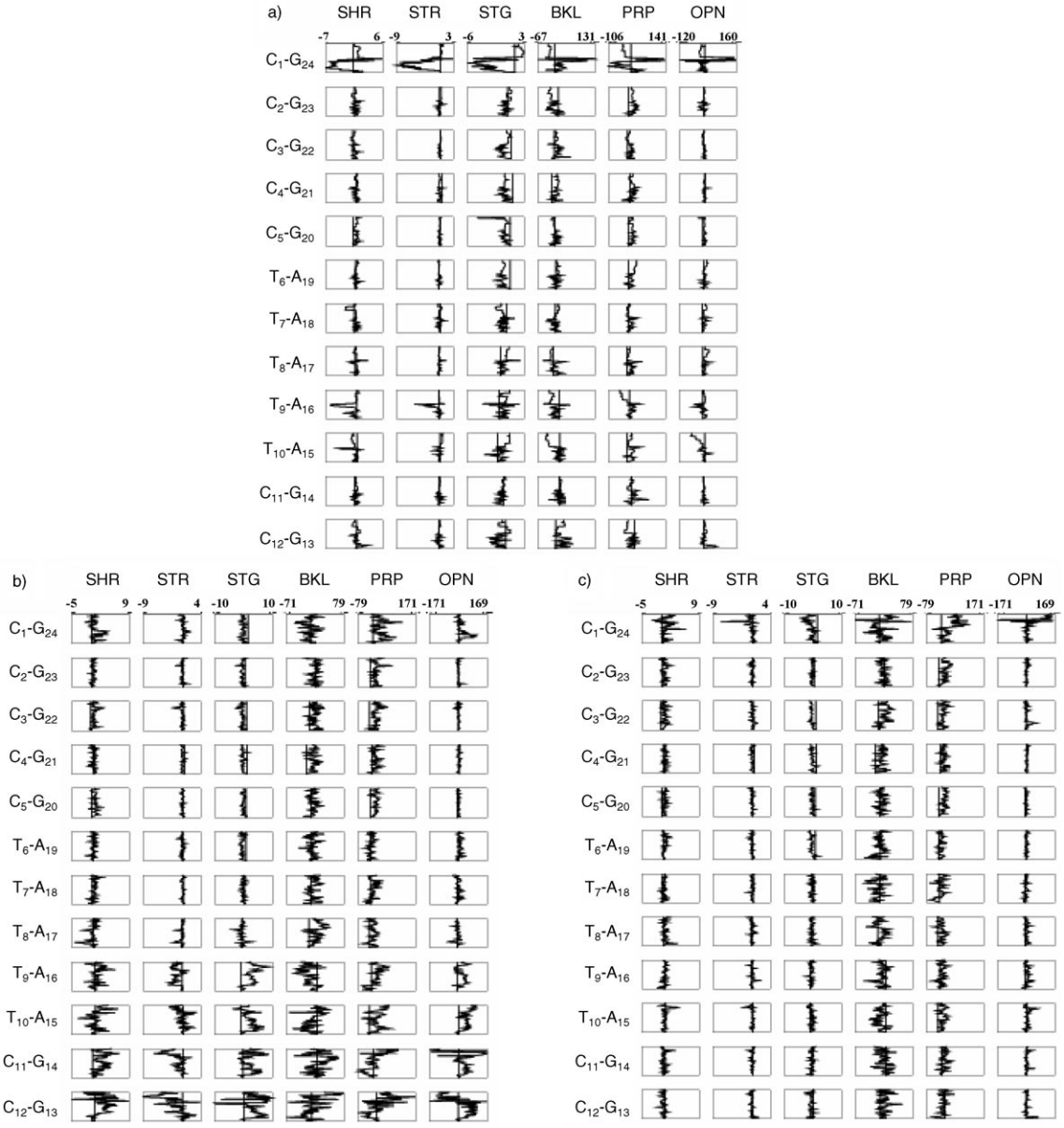


Figure 3. Helicoidal parameters for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^{0}. Time is given on the vertical axis (0–2 ns, 0 top) and the helicoidal parameter on the horizontal axis (scale as marked). For a definition of these parameters[38] see Supporting Information.

simply inconsistent with a stable Watson–Crick base pair: shear deformations vary by as much as 8 Å during a simulation, while the base pair "stretches" are, at times, 3–4 Å *less* than their equilibrium value.

Visual inspection of the DNA/C^{4+} trajectory showed that this apparent disruption of the double helix below base-pair T9–A16 was actually a complex rearrangement amongst the base pairs, resulting in the formation of a new pairing between T9 and A15, so that T10 and A16 were left unpaired (see Figure 4). The sequence of events for the formation of this defect was as follows:

1) Some stress became evident in base pairs 10–12 from about 170 ps:
2) T9–A16 broke apart at 390 ps;
3) a new base pair, T9–A15, formed within the following 10 ps.

T10 was then observed to swing out of the helix, kinking the backbone, and subsequently disrupting the last two base pairs (C11–G14 and C12–G13). It is interesting to note that this activity occurred adjacent to, rather than at, the cylinder binding site: the cylinder remained bound to the base pairs numbered 5–8 (i.e., C5–G20 to T8–A17) throughout the simulation.

This change observed in the base-pair stacking is also revealed in the pattern of hydrogen bonding between the nucleic acids at the end of the simulation. Hydrogen-bond

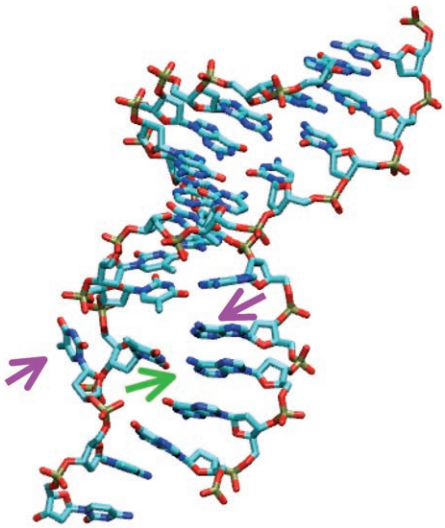


Figure 4. Snapshot of DNA showing the base-pair defect induced by C^{4+}. The green arrow points to the slipped mismatch between T9–A15; the magenta arrows indicate the resulting unmatched bases, T10 and A16.

lengths were obtained from the average DNA structure, as calculated by the 3DNA package, using the final 50 ps of each trajectory. This time window was found to be long enough to smooth out the instantaneous vibrations, but still short enough to avoid artefacts due to long timescale conformational motions of the DNA. The results for the two DNA/cylinder simulations are listed in Table 1 and confirm the formation of a mismatched Watson–Crick base pair T9–A15, with the two hydrogen bonds, in the presence of C^{4+}.

Dial plots of the DNA backbone torsion angles from the three sets of simulations are given in Figure 5. Data for the uncomplexed DNA are consistent with the average values expected for B-DNA, and show no evidence of a transition to one of the other forms. The presence of the cylinder does change the DNA backbone conformation, but the changes are smaller than was seen in the helicoidal parameters. Interestingly, while the helicoidal parameters indicated that the base-pair structure was more constrained in DNA/C^0 than in the uncomplexed DNA, the opposite appears to be true for the backbone, with the backbone torsions showing larger fluctuations in the presence of C^0 than without. The changes in base-pair stacking induced by C^{4+} are also seen in the backbone, but the difference between the two cylinders is less obvious than it was with the direct measures of base-pair geometry. The most obvious indication is in the C3′-O-P-O (ζ) and O-P-O-C5′ (α) angles, which show almost random angular variations between base pairs 10–12 in the presence of C^{4+}, suggestive of rotation that is unhindered by base-pair formation. This effect appears to be more pronounced at the end of the first strand (nucleotides T10 to C12) than on the complementary strand (nucleotides G13 to A15).

A schematic representation of the global helix axis after 2 ns is given in Figure 6. We note that quantifying curvature in the DNA helix can be problematical. Any measure must be able to distinguish between the local oscillations in base-pair orientation found in, particularly, the A-form, and a real bend in the overall helix direction. This can be particularly difficult in short strands of DNA, such as the dodecamer studied here, since one cannot use the behaviour of the helix beyond the binding site to confirm the persistence of any bend. To measure the extent of bending we have used the 3DNA definition of the local helix direction. This uses the geometry of any two adjacent base pairs (a step) to define a unit vector, $\boldsymbol{U}_i$, along the helix axis at that step. Bending of the DNA can then be monitored by defining an angle, $\vartheta_i = \cos^{-1}(\boldsymbol{U}_{ref} \cdot \boldsymbol{U}_i)$, which describes the total bend in the DNA between some reference step (taken to be the first stable step in this work) and the *i*th base-pair step. Most importantly, this definition gives perfectly aligned $\boldsymbol{U}_i$ ($\vartheta_i = 0$ for every *i*) in *both* the canonical A and B forms, so that deviations from 0 can be interpreted as bending of the DNA.

ϑ_i values for the, averaged, DNA structures after 2 ns are shown in Figure 7; we reiterate that the DNA response was complete on this timescale, and no further relaxation was observed in simulations continued to 5 ns. Thus, these plots give a good indication of both the global bend induced in the DNA by the ligands, and how it is distributed along the dodecamer. Both C^{4+} and C^0 can be seen to bend the DNA by about 40°, which compares very favourably with the 40–60° per cylinder measured experimentally. In the presence of C^{4+} this bend is strongly localised on steps 5 and 6 (base pairs 5–7), but with the consequent degradation of base pairs 9–12 noted above. In contrast, the

Table 1. Intra-base pair hydrogen bond length, taken from the average DNA structure observed during the final 50 ps of the simulation; blank values indicate no hydrogen bond was found. The mismatch induced by C^{4+} is indicated in bold.

Base pairings	Length [Å]					
	DNA/C^{4+}			DNA/C^0		
	O2–N2	N3–N1	N4–O6	O2–N2	N3–N1	N4–O6
C1–G24	2.75	3.06	3.22			
C2–G23	2.96	3.04	2.97	2.82	3.00	3.02
C3–G22	2.86	3.00	3.01	2.86	3.01	2.97
C4–G21	2.80	3.03	3.10	2.80	3.15	3.35
C5–G20	2.91	3.05	2.99	3.03	2.99	2.84
T6–A19		2.98	3.18		2.99	3.01
T7–A18		3.05	2.91		3.06	2.95
T8–A17		3.03	2.89		3.00	2.87
T9–A16					3.03	3.00
T9–A15		**3.17**	**2.89**			
T10–A15					3.13	2.88
C11–G14				3.00	3.01	2.94
C12–G13				2.85	3.01	3.00

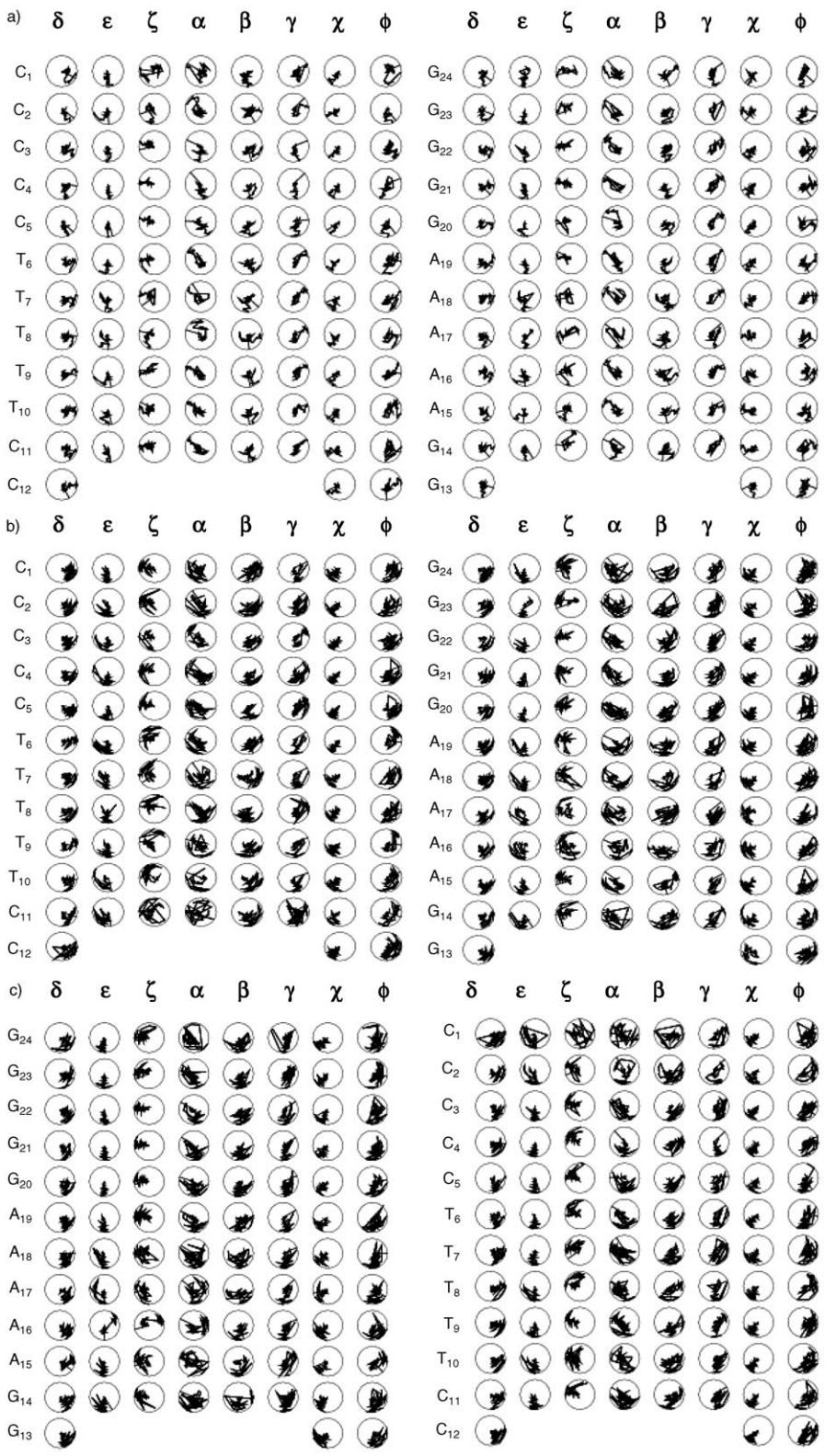


Figure 5. DNA backbone torsion angles for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^{0}. Circular polar plot of time (radius, 0–2 ns) against torsion angle (0° top, 90° right).

40° bend for C^{0} is spread over 10 base pairs, which probably explains why the integrity of the duplex is retained in this case. The uncomplexed DNA shows a net bend of only 10–12° over 10 base pairs, with some evidence of more pronounced curvature over the first few steps.

Another common descriptor for helix curvature is the normal vector plot (NVPs).[43] NVPs are projections of the unit vector normal to the plane of each base pair onto a plane that is normal to the helix axis. Linear segments of the DNA can then be identified as clusters of neighbouring points, while coiling is seen as a steady change of orientation across a number of adjacent base pairs. NVPs calculated at the end of each simulation are presented in Figure 8. The uncomplexed DNA shows a reasonably tight cluster of points. In the presence of the neutralised cylinder (C^{0}), however, there is a large and systematic variation in orientation across base pairs 3–9, consistent with the strong curvature noted above. For the fully charged cylinder, C^{4+}, there is strong curvature but as was seen with the cumulative bending angle (Figure 7), this effect is localised between base pairs 6 and 7, that is, to the start of the region in which the ligand is bound, and is suggestive of a localised kink in the double helix.

Various parameters describing the geometry of base-pair steps have also been calculated using 3DNA, and the average structure obtained during the final 50 ps portion of the trajectory. In Figure 9 we report values of the slide displacement, and roll, twist and tilt angles,[40,41] for all stable base pairs. There are several important points to note from these plots. In the first place, the step parameters observed in the presence of the cylinders are not readily identified with any of the common DNA conformations. Slide values below −1 Å, as observed along most of the stable duplex in the presence of both C^{4+} and C^{0}, are often taken to be indica-

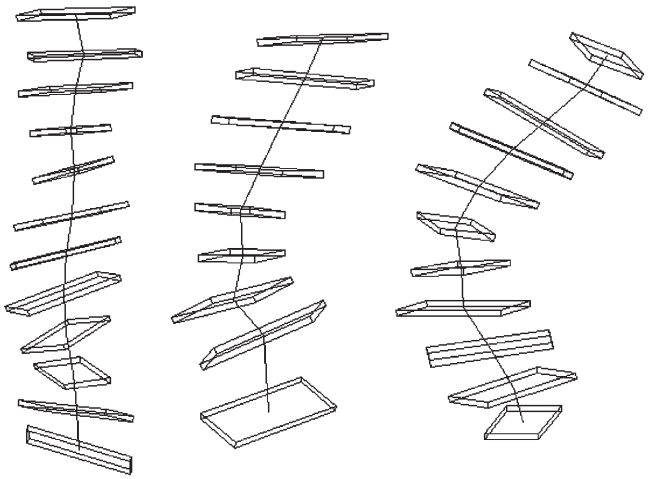


Figure 6. Representation of the average DNA structure in the uncomplexed (left), DNA/C^{4+} (middle) and DNA/C^0 (right) simulations. Averages were calculated from the final 50 ps of each simulation using 3DNA.[42] Only the stable base pairs are depicted.

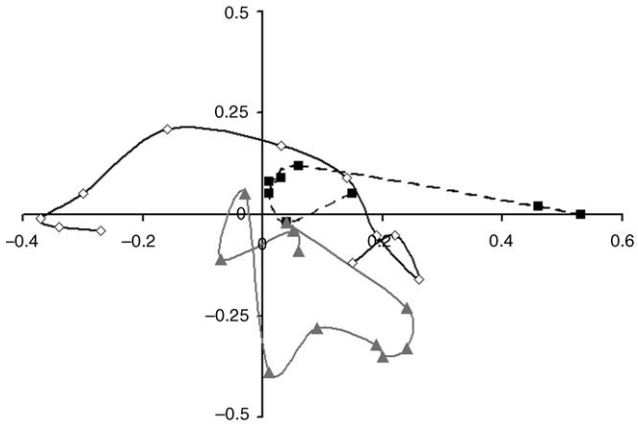


Figure 8. Normal vector plots (NVPs) for the average DNA structure calculated from the final 50 ps of each simulation using the CHARMM22 force field. NVPs are a projection of a unit vector normal to each base pair onto a plane that is normal to the local helical axis, with the points corresponding to successive base pairs being joined by a smooth line. ◇: C^0, ■: C^{4+}, ▲: uncomplexed.

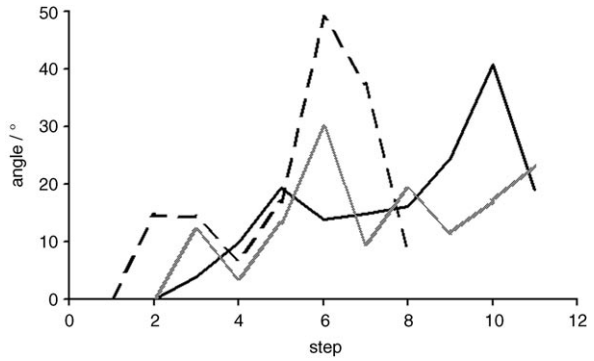


Figure 7. Degree of bend in the DNA, calculated from the local helix direction as defined by 3DNA. The angles are defined as $\vartheta_i = \cos^{-1}(\boldsymbol{U}_{ref} \cdot \boldsymbol{U}_i)$, where $\boldsymbol{U}_{ref}$ is the helix director for the first stable step. The helix bend (left) is identically zero for both the canonical A- and B-forms. ——: C^0, -----: C^{4+}, ——: uncomplexed.

tive of A-DNA. However, both the slide and roll values obtained in this work are really intermediate between the values expected for A- and B-DNA, while, for C^0, the twist is more consistent with B-DNA. On balance, 3DNA assigns the duplex as having the B-form in the presence of the neutral cylinder, and fails to make any assignment in the presence of C^{4+}. It is also interesting to note that both the slide and the roll angles take their most extreme values at step 6, which is both the start of the region to which the cylinders bound and the start of the A-T tract. Finally, we note that the twist angles indicate that neither C^{4+} nor C^0 cause significant winding or unwinding of the DNA—either locally or globally. The average twist for C^0 is 33.4° per step, which is intermediate between the A- and B-forms. A slightly lower value (29.4° per step) is seen in the presence of C^{4+}, but the extent to which this is influenced by the mismatched base pairing is unclear. This behaviour is in contrast to some proteins, such as 434 Cro, which bend the DNA via alternating regions of under- and overwinding.[1]

A concomitant question when DNA bending is observed is to ask how the size of major and minor grooves is affected. Following Hassan and Calladine,[48] we have monitored these by measuring the distance between appropriately displaced phosphate groups, with the P atom representing the position of the phosphate. The width of the major groove can be defined as the distance between the P atom in the *n*th nucleotide, counting from the 5′ end of one strand, and the P atom of the *n*+3rd nucleotide from the 3′ end of the complementary strand; this numbering assumes 5′ nucleotides. Similarly, for the minor groove one can use the distance between the P atoms in the *n*th nucleotide from the 5′-end of one strand with the *n*−4th nucleotide from the 3′-end of the complementary strand. The distribution of these inter-phosphate distances was determined by analyzing every configuration saved from the final 0.25 ns of each simulation. To avoid artefacts from end-fraying, those phosphates that were adjacent to the first and last base pair were omitted from the analysis. This left five measurements of the major groove width, and six for the minor groove. It should be noted that, for a dodecamer, this meant that the minor and major groove measurements were made on largely different regions of the DNA, with only two base pairs (T6–A19 and T7–A18) being used to measure both the major and the minor groove widths.

Unlike crystal structure analyses, the DNA shows considerable flexibility during a simulation, and so gives rise to a distribution of groove widths for each pair of phosphate groups. The average groove width and its standard deviation are presented in Figure 10. It is striking that the DNA-coiling induced by C^0 does not in any way affect the width of the major groove—the site into which C^0 binds. An enlargement in the width of the minor groove (ca. 2.5 Å) is seen in the vicinity of the C^0 binding site, but only for four of the

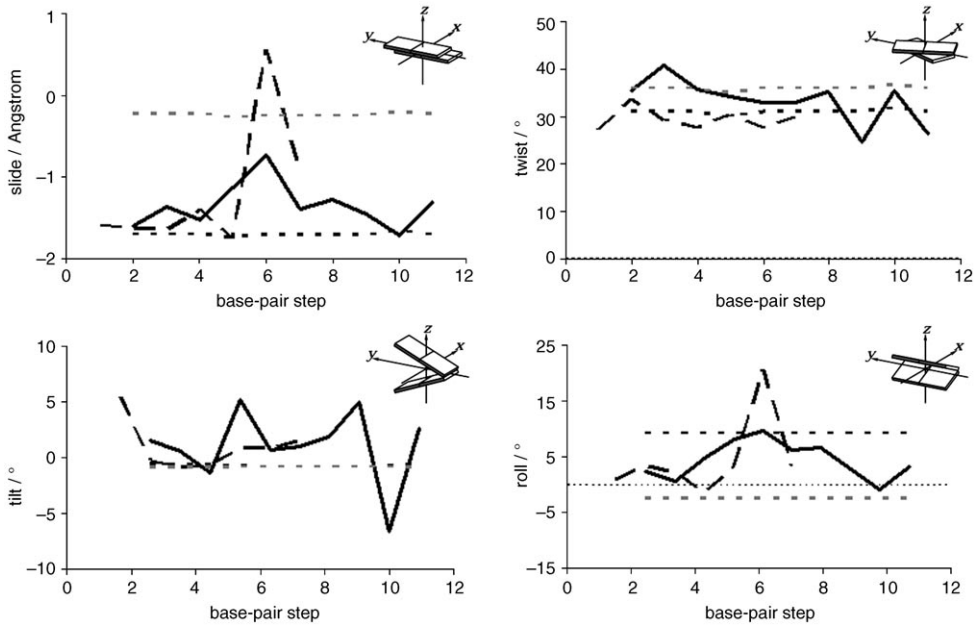


Figure 9. Slide displacement, and twist, tilt and roll angles between adjacent base pairs; a schematic definition of each inter-base pair coordinate is given with the plot. The horizontal axis corresponds to the step between base pairs, so that step 1 compares the first two base pairs (C1–G24 and C2–G23), etc. Only stable base-pairs have been included. Data for the canonical A- (-----) and B-forms (- - - - -) are given as a reference; ——: C^0, – – –: C^{4+}.

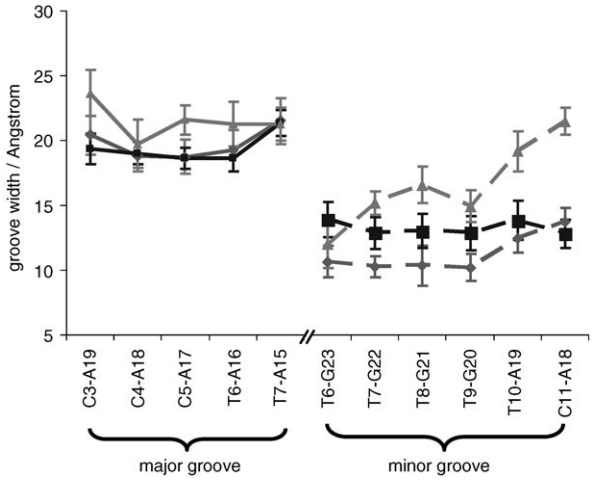


Figure 10. Widths of the major (solid) and minor (dashed) grooves, as determined from specified inter-phosphate distances. Calculations were performed with CHARMM22. "Error bars" denote the width of the distribution of distances (± one standard deviation), not the uncertainty in the calculation. No correction has been made for the size of the phosphate group; ◆: uncomplexed, ■: C^0, ▲: C^{4+}.

six inter-phosphate distances found in this region (those involving nucleosides T6 through T9, but not A18 and A19).

In contrast to C^0, C^{4+} has a substantial effect on the widths of both the major and minor groove. The major groove width increases by an average of 1.8 Å compared with both the uncomplexed DNA and the DNA with bound C^0. Even larger increases are seen in the width of the minor groove. However the largest increases (nearly 8 Å for C11–A18) are in regions where the backbone torsion angles α and ζ have already been shown to be ill defined (Figure 5) and so probably indicate that there is no well defined minor groove in this region when C^{4+} is bound to the DNA.

In summary, the supramolecular cylinder, **1**, is seen to have a strong influence on the conformation of DNA as modelled with the CHARMM22 force field. The interactions within a neutralised analogue of **1** were sufficient to induce a bend of about 40° in the DNA helix axis, which is very similar to the angle seen in experiments with **1**. When the tetracationic charge is also accounted for, **1** was found to strain the base pairing within the duplex, and induced a mispairing of A–T bases adjacent to the ligand binding site. It is important to stress that the induction of base-pair defects was reproducible. Four separate C^{4+} simulations were performed, starting with different low-energy cylinder/DNA binding sites and with different arrangements of the solvent and counterions. In each case a mismatched A–T base pair formed within the first 0.5 ns. It is possible that the defects were introduced as part of the coiling process (i.e., whether the additional strain during coiling made the DNA susceptible to defects in the base-pair sequence) in which case they would not reflect a global free energy minimum. To test this, a further simulation was performed starting from a stable configuration obtained after 2 ns of the DNA/C^0 simulation and changing the atomic charges back to those of C^{4+}; four Na^+ ions were also removed to maintain electrical neutrality in the system. An equivalent extended simulation with C^0 was used as a control. Within just 0.25 ns, mispairings in the A–T region again appeared in the presence of C^{4+} (Supporting Information, Figure 4), but no mispairings appeared with

C^0 during an additional 3 ns. We conclude that strong coiling of DNA by **1**, coupled with a disruption of the duplex base pairing, is a real prediction from the CHARMM22 force field.

Force-field dependence: The two DNA force fields have been compared by measuring the extent to which a stable binding geometry obtained with one force field is maintained by the other. This amounts to determining whether an important minimum on the free energy surface of one force field is reproduced with the other. By starting in a well-defined free-energy minimum, instead of repeating the docking protocol, one avoids the possibility of simulations with different potentials simply converging on different local minima and thereby ends up with an unambiguous comparison of the two force fields. Accordingly, simulations with the CHARMM27 force field were started from the stable conformation obtained after 2 ns in the C^0/CHARMM22 simulation.

Simulations were again performed with both tetracationic and neutralised cylinders and are denoted ${}_{27}C^{4+}$ and ${}_{27}C^0$, respectively. For the ${}_{27}C^{4+}$ system, four Na^+ ions chosen at random were removed to maintain electrical neutrality. The system was then relaxed in a series of short, restrained simulations, as outlined in Section Computational Methods. All restraints (i.e., tethering potentials) were then removed and the system simulated for an additional 1.5 ns. The 3DNA analysis of bend and step parameters, obtained from the average structure at the end of these simulations, is presented in Figures 10–12, while the distribution of groove widths observed during the final 0.25 ns is depicted in Figure 14. Where clarity permits, results for CHARMM22 are also depicted for comparison.

In general, the CHARMM27 DNA is less strongly influenced by the cylinder than is the CHARMM22 model. For both ${}_{27}C^{4+}$ and ${}_{27}C^0$, the DNA uncoils slightly to give an overall bend of only about 20° (Figure 11), while the normal vector plots (Figure 12) remain relatively compact. At the same time the slide parameters (Figure 13) indicate a shift to the B-form that is known to be favoured by CHARMM27 for an aqueous duplex in the absence of a ligand. No evidence was found that either cylinder disrupted the base pairings at any stage during the simulations, with the average structure from the end of the simulations showing the expected three (G–C) and two (A–T) hydrogen bonds for every base pair; at least one of the hydrogen-bond lengths was less than 3.0 Å in every base pair, while none was more than 3.13 Å. This is a major difference from the analogous CHARMM22 simulation described above, where reinstating the $+4\,e$ cylinder charge on the DNA/C^0 complex induced base-pair mismatches within just 0.25 ns.

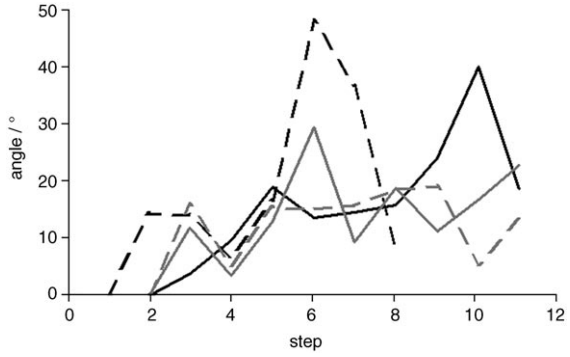


Figure 11. Global bend of the DNA, calculated as for Figure 6. CHARMM27 results are given as dashed lines, and CHARMM22 as the solid lines; ——: C^0, -----: C^{4+}, ——: ${}_{27}C^0$, - - - - -: ${}_{27}C^{4+}$.

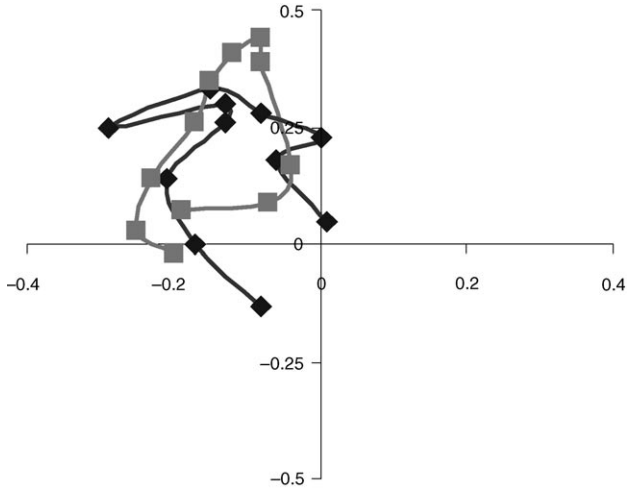


Figure 12. Normal vector plots for the average DNA structure obtained during the final 50 ps of the CHARMM27 simulations. ◆: ${}_{27}C^0$, ■: ${}_{27}C^{4+}$.

Perhaps most surprisingly, the cylinder charge appears to have very little effect on the CHARMM27 DNA. The normal vector plots, global bend and step parameters (with the exception of the roll angle at the 6th step) all show remarkably similar behaviour for the ${}_{27}C^{4+}$ and ${}_{27}C^0$ systems. Given the strongly ionic character of DNA and its flexibility, this apparent invariance to the ligand charge is unexpected. One property that does show differences between the two force fields is the groove widths. Relative to CHARMM22/C^0, the major groove of CHARMM27-DNA contracts in the presence of C^{4+}, but expands in the presence of C^0. We note that CHARMM27 uses a more strongly charged DNA backbone than does CHARMM22, with the PO_4 net charge being $-1.2\,e$, compared with $-1.0\,e$ for CHARMM22. This increased negative charge may lead the DNA backbone to contract more strongly onto the tetracationic cylinder, resulting in a stiffening and straightening of the DNA backbone and thereby inhibiting the subsequent bending of the DNA. A similar response to ligand charge has been seen with netropsin bound in the minor groove,[24] but greater scope for DNA relaxation might have been anticipated when the ligand binds in the major groove.

The comparison between these two force fields highlights several influences on the DNA response. In the first place we note that both force fields predict the neutralised ligand to coil DNA at least as strongly as the tetra cationic ligand. Previous work has shown that regions of low dielectric con-

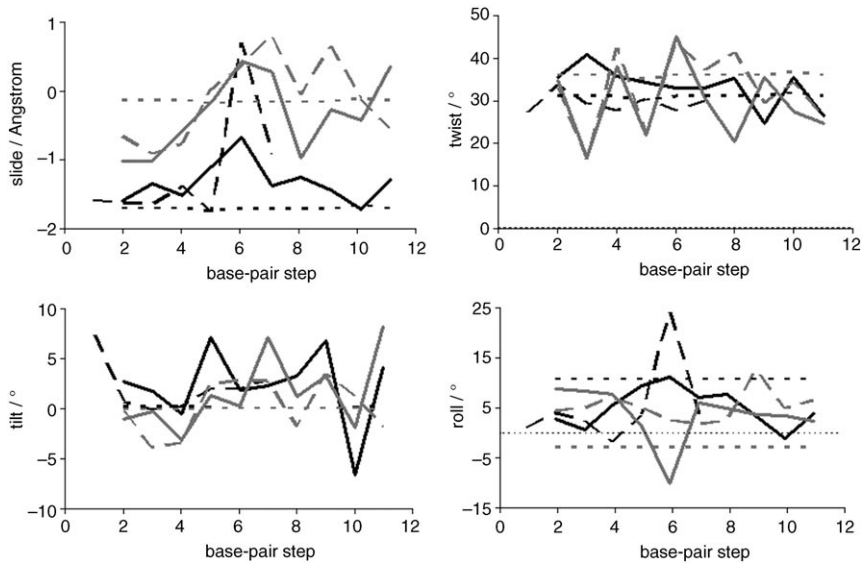


Figure 13. Selected step parameters for stable base pairs, calculated as for Figure 8. Values for the canonical A- and B-forms are given for reference, ——: C^0, -----: C^{4+}, ——: $_{27}C^0$, ------: $_{27}C^{4+}$, •••••: A-form, ·····: B-form.

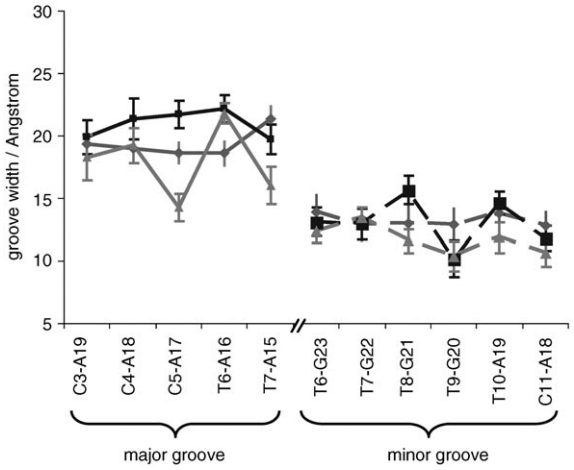


Figure 14. Groove widths calculated during the final 0.25 ns of the CHARMM27 simulations. Solid lines indicate the major groove and dashed lines the minor groove,◆: C^0, ■:$_{27}C^0$, ▲: $_{27}C^{4+}$.

stant enhance the phosphate repulsions in DNA and cause it to bend *away* from the low dielectric region.[49] Although the neutralised ligand does generate electric field gradients, its enhanced negative surface charge should also accentuate the phosphate repulsions and induce the DNA to bend away from the ligand, or at least coil less tightly about it. The fact that this is not observed, and indeed the coiling may be enhanced for the neutralised ligand, implies that it is the other forces—the van der Waals forces—that drive the coiling.

Ligand charge does have some affect on the DNA response, but its principal manifestation is different in the two force fields. The contrast is instructive. In CHARMM27, the $+4\,e$ charge attracts the negatively charged DNA backbone, causing the major groove to contract onto the cylinder. With CHARMM22 its effect is to strain the hydrogen bonding between base pairs in the duplex. The induction of such strain is reasonable given that most force fields describe hydrogen bonding primarily through electrostatic interactions. Orbital interactions, such as charge transfer and π-electron polarisation, do contribute strongly to the hydrogen bonding between base pairs,[50,51] but these are also likely to be strongly perturbed by the proximity of a tetracationic ligand. In reality, both the strain to the base pair hydrogen bonds and the contraction of the DNA backbone onto the ligand are likely to be present, and getting the balance right will be an important validation of the force field. In this context, the CHARMM22 result that duplex base-pairing will be disrupted by C^{4+} provides an interesting prediction that awaits the measurement of definitive X-ray or NMR structures.

Sensitivity to cylinder charge: If, as indicated above, the disruption of base pairs in a DNA duplex with CHARMM22 is due to a competition between inter-base pair hydrogen bonding and ligand base electrostatics, then it is of interest to identify how strong the ligand electrostatics need to be in order to effect the base pair mismatches. Analogous simulations to those reported in the last section were performed with CHARMM22 and with a variant of the cylinder in which the overall charge was set to $+3\,e$ (C^{3+}), again achieved by a constant shift of the atomic charges for all C and non-polar H atoms. Within 0.25 ns this system also showed disruption of the base pairing within the double helix (see Figure 15). Similar calculations with a $+2\,e$ variant of the cylinder (not shown) gave results that were intermediate between C^{3+} and C^0, with frequent bifurcated pairings in which an A was found to bridge two T bases in the opposite strand, but these did not lead to irreversible breakdown in the base pairing on a nanosecond timescale. We conclude that some disruption of the duplex base pairs is a robust feature of the CHARMM22 model, and is not especially sensitive to the parameterisation of the cylinder charge distribution.

Conclusion

Our molecular dynamics simulations of the supramolecular cylinder, **1**, in explicit solvent show major groove binding of the cylinder on the DNA and coiling of the DNA in re-

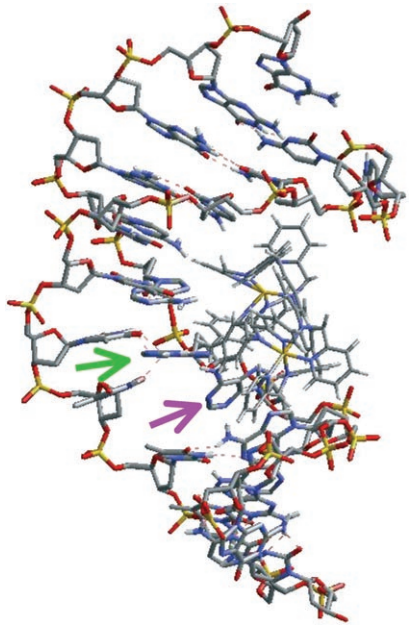


Figure 15. Configuration of DNA 0.25 ns after continuing from the final configuration of the 2 ns DNA/C^0 simulation with a $+3\,e$ cylinder (C^{3+}). The magenta arrow points to an unpaired T, while the green arrow points to two A bases paired to a single T.

sponse to this binding. This is consistent with all available experimental observations. Moreover, the simulations have, as envisaged, yielded molecular level information about the coiling effect. Other highly charged cations, such as spermine and cobalt(III)–hexamine,[17,22,52,53] bend DNA by a few degrees per ligand and cause it to condense into poly-molecular aggregates (often of well-defined shape). Our simulations indicate that the DNA bend induced by **1** is an order of magnitude greater than this: the simulations give 20–40°, depending on the force-field model, which compares favourably with the experimental measurements of 40–60° per ligand. Somewhat surprisingly, the tetracationic charge of the supramolecular cylinder was found *not* to be essential for inducing the coiling, as a neutral analogue was also found to bend the DNA. Instead the molecular shape appears to be the key factor and must give rise to a coil with a pitch which ensures an intramolecular effect observed experimentally with this ligand, rather than the more commonly observed intermolecular aggregation. This highlights the significance of our molecular design which creates, through metallo-supramolecular assembly, a cylindrical agent of very similar shape and dimensions to the various cylindrical motifs employed by proteins for DNA recognition and structure control.

This is *not* to say electrostatic interactions were unimportant. Our simulations identified two competing influences arising from a cationic ligand in the major groove that are likely to be general effects. On the one hand, the large cationic charge attracted the phosphate backbone, causing the major groove to contract around the ligand. On the other hand, such a large positive charge density close to the base pairs places strain upon the inter-base hydrogen bonding that holds the DNA duplex together. The balance of these two effects did depend on the choice of DNA force field, but the latter effect was sufficiently large that C^{4+} was observed to generate base-pair mismatches in some systems. When these mismatches were observed, they occurred in the A–T tract and adjacent to, rather than at, the site at which the cylinder bound. Any weakening of the base pair hydrogen bonding at or near this site could prime the DNA for transcription or replication processes. This is an unexpected molecular-level feature associated with the experimentally observed cylinder binding and DNA conformational change. Intriguingly, we note that many proteins also bend DNA at locations adjacent to the site of transcription. For example, the CAP (catabolic gene activating protein) binds adjacent to the initial site at which RNA polymerase acts and causes a 50-fold increase in the rate of transcription initiation.[54] This activation is usually ascribed to the provision of an interaction surface for the polymerase, thereby enhancing binding to the promoter site. Our C^{4+}/DNA simulations suggest there may be an alternative mechanism: that the bending induced by CAP binding can help to open the DNA at an adjacent site, and could thus contribute to the increased rate at which transcription is initiated in the presence of promoters. This opens up new potential applications for these DNA-coiling metallo-supramolecular cylinders.

Thus we conclude that the tetracationic supramolecular cylinder, **1**, does have two substantive effects on DNA. In the first place it can induce coiling of the DNA, with potentially very large curvature being introduced in DNA dodecamers. However, at the same time it weakens the hydrogen bonding between base pairs near the binding site and so can tend to introduce defects into the alignment of the base-pair sequences in the two strands of the double helix. In our simulations this occurred consistently within the A–T rich region of the DNA and resulted in misaligned A–T cross-links between the two strands. It is not clear whether this represents a selectivity for A–T over G–C and simulations with other DNA sequences are in progress to resolve this issue.[1,2]

Acknowledgements

We thank the EPSRC (S.K.; postgraduate studentship) and the European Union (MARCY RTN; HPRN-CT-2002-00175) for support. M.J.H. is the Royal Society of Chemistry Sir Edward Frankland Fellow 2004-5.

[1] See, for example, J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th ed., Freeman, New York, **2002**.
[2] J. P. Baak, F. R. Path, M. A. Hermsen, G. Meijer, J. Schmidt, E. A. Janssen, *Eur. J. Cancer* **2003**, *39*, 1199–1215.
[3] C. F. Calkhoven, C. Muller, A. Leutz, *Trends Mol. Med.* **2002**, *8*, 577–583.
[4] L. Rogge, *Arthritis Res. Ther.* **2003**, *5*, 47–50.
[5] B. S. Parekh, G. W. Hatfield, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1173–1177.

[6] R. R. Sinden, *DNA Structure and Function*, Academic Press, London, **1994**.
[7] R. E. Dickerson, *Nucleic Acids Res.* **1998**, *26*, 1906–1926.
[8] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland, New York, 2nd ed., **1999**.
[9] B. C. Baguley, *Mol. Cell. Biochem.* **1982**, *43*, 167–181.
[10] W. A. Denny, *Anti-Cancer Drug Des.* **1989**, *4*, 241–263.
[11] S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA, **1994**.
[12] D. Z. Coggan, I. S. Haworth, P. J. Bates, A. Robinson, A. Rodger, *Inorg. Chem.* **1999**, *38*, 4486–4497.
[13] B. Onfelt, P. Lincoln, B. Nordén, *J. Am. Chem. Soc.* **1999**, *121*, 10846–10847.
[14] K. E. Erkkila, D. T. Odom, J. K. Barton, *Chem. Rev.* **1999**, *99*, 2777–2795.
[15] M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 879–884.
[16] I. Meistermann, A. Rodger, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, S. Khalid, P. M. Rodger, J. Peberdy, C. J. Isaac, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074.
[17] A. Rodger, K. J. Sanders, M. J. Hannon, I. Meistermann, A. Parkinson, D. S. Vidler, I. S. Haworth, *Chirality* **2000**, *12*, 221–236.
[18] G. S. Manning, K. K. Ebralidse A. D. Mirzabekov, A. Rich, *J. Biomol. Struct. Dyn.* **1989**, *6*, 877–889.
[19] E. Moldrheim, M. J. Hannon, I. Meistermann, A. Rodger, E. Sletten, *J. Biol. Inorg. Chem.* **2002**, *7*, 781–790.
[20] I. S. Haworth, A. Elcock, A. Rodger, W. G. Richards, *J. Biomol. Struct. Dyn.* **1991**, *9*, 553–569.
[21] B. P. Hudson, J. K. Barton, *J. Am. Chem. Soc.* **1998**, *120*, 6877–6888.
[22] T. E. Cheatham, P. A. Kollman, *Structure* **1997**, *5*, 1297–1311.
[23] Y. Y. Fang, B. D. Ray, C. A. Claussen, K. B. Lipkowitz, E. C. Long, *J. Am. Chem. Soc.* **2004**, *126*, 5403–5412.
[24] B. Wellenzohn, W. Flader, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *Biopolymers* **2001**, *61*, 276–286.
[25] J. Dolenc, C. Oostenbrink, J. Koller, W. F. van Gunsteren, *Nucleic Acids Res.* **2005**, *33*, 725–733.
[26] A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnik, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
[27] N. Foloppe, A. D. MacKerell, Jr., *J. Comput. Chem.* **2000**, *21* 86–104.
[28] S. Y. Reddy, F. Leclerc, M. Karplus, *Biophys. J.* **2003**, *84*, 1421–1449.
[29] W. Smith, T. R. Forester, *J. Mol. Graphics* **1996**, *14*,136.
[30] a) M. A. San Miguel, R. Marrington, P. M. Rodger, A. Rodger, C. Robinson, *Eur. J. Biochem.* **2003**, *270*, 3345–3352; b) M. A. San Miguel, P. M. Rodger, *Phys. Chem. Chem. Phys.* **2003**, *5*, 575–581; c) R. Lukac, A. J. Clark, M. A. San Miguel, A. Rodger, P. M. Rodger, *J. Mol. Liq.* **2002**, *101*, 261–272; d) T. Astley, G. G. Birch, M. G. B. Drew, P. M. Rodger, *J. Phys. Chem. A* **1999**, *103*, 5080–5090.
[31] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
[32] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comp. Physiol.* **1977**, *23*, 327.
[33] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford (UK), **1987**.
[34] M. Feig, B. M. Pettitt, *Biophys. J.* **1998**, *75*, 134–149.
[35] N. Korolev, A. P. Lyubartsev, A. Laaksonen, L. Nordenskiöld, *Biophys. J.* **2002**, *82*, 2860–2875.
[36] S. Khalid, PhD Thesis, University of Warwick (UK), **2004**.
[37] B. Wellenzohn, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *J. Am. Chem. Soc.* **2000**, *122*, 3927–3931.
[38] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 655–667.
[39] G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 669–699.
[40] R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. M. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C. S. Tung, A. H. J. Wang, V. B. Zhurkin, *EMBO. J.* **1989**, *8*, 1–4.
[41] W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinmann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, H. M. Berman, *J. Mol. Biol.* **2001**, *313*, 229–237.
[42] X. J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
[43] D. L. Beveridge, S. B. Dixit, G. Barreiro, K. M. Thayer, *Biopolymers* **2004**, *73*, 380–403; D. McDonald, P. Lu, *Curr. Opin. Struct. Biol.* **2002**, *12*, 337–343.
[44] R. E. Dickerson, T. K. Chiu, *Biopolymers* **1997**, *44*, 361–403.
[45] R. R. Sinden, *DNA Structure and Function*, Academic Press, New York, **1994**.
[46] A. Barbic, D. P. Zimmer, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2369–2373.
[47] J. Hizver, H. Rozenberg, F. Frowlow, D. Rabinovich, Z. Shakked, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8490–8495.
[48] M. A. Hassan, C. R. Calladine, *J. Mol. Biol.* **1998**, *282*, 331–343.
[49] A. H. Elcock, A. McCammon, *J. Am. Chem. Soc.* **1996**, *118*, 3787–3788.
[50] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **2002**, *114*, 2194–2197; *Angew. Chem. Int. Ed.* **2002**, *41*, 2092–2095.
[51] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111*, 3120–3122; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942–2945.
[52] I. S. Haworth, A. Rodger, W. G. Richards, *Proc. R. Soc. London Ser. B* **1991**, *244*, 107–116.
[53] K. Bryson, R. J. Greenall, *J. Biomol. Struct. Dyn.* **2000**, *18*, 393–412.
[54] See reference [1], p. 873.

Received: September 22, 2005
Published online: February 23, 2006

DOI: 10.1002/chem.200600060

A DNA-Binding Copper(I) Metallosupramolecular Cylinder that Acts as an Artificial Nuclease

Laura J. Childs,[b] Jaroslav Malina,[b] Britt Elin Rolfsnes,[d] Mirela Pascu,[a, b] Maria J. Prieto,[c] Mark J. Broome,[b] P. Mark Rodger,[b] Einar Sletten,[d] Virtudes Moreno,*[c] Alison Rodger,*[b] and Michael J. Hannon*[a, b]

Abstract: The DNA binding of a dicationic pyridylimine-based dicopper(I) metallosupramolecular cylinder is reported together with its ability to act as an artificial nuclease. The cylinder binds strongly to DNA; more strongly than the spherical dication $[Ru(phen)_3]^{2+}$ (phen = 1,10-phenanthroline), but more weakly than the corresponding tetracationic cylinders. DNA coiling effects are not observed with this dication, in contrast to the situation with the previously reported tetracationic cylinder involving a similar ligand. Linear dichroism (LD) data suggests that the dicopper cylinder binds in a different orientation from that of the tetracationic iron cylinder. Furthermore, the dicopper cylinder shows DNA-cleavage activity in the presence of peroxide. Of particular note is that the cylinder displays a marked and unusual ability to cleave both DNA strands at the same site, probably reflecting its dinuclear nature and possibly its mode of binding to the DNA.

Keywords: artificial nucleases · bioinorganic chemistry · copper · DNA cleavage · supramolecular chemistry

Introduction

The genetic "code" of most living organisms is contained in their DNA. In mammalian cells, approximately 3 cm of DNA carries the information needed to assemble and sustain the entire organism.[1] Processing of the genetic information (transcription or replication) requires sequence-specific recognition. This is achieved by proteins, most frequently through major-groove binding interactions, as the size and shape of the major groove of B-DNA varies more with base sequence. In addition, this groove contains more hydrogen-bond donor and acceptor sites whose spatial dispositions are sequence dependent. For example, zinc-finger subunits are found in many DNA-binding proteins and contain a zinc ion that holds together a cylindrical loop of amino acids that recognises the DNA major groove.[1,2] Other helical DNA-recognition motifs employed by proteins include α-helices, and such protein recognition of DNA usually occurs through noncovalent interactions. Traditional *synthetic* DNA-recognition agents are generally much smaller than those of proteins and, as a consequence, frequently act through either intercalation or minor-groove binding.[3,4] Synthetic agents that target the DNA major groove by binding through noncovalent interactions have the potential to be powerful agents for sequence-specific DNA recognition. We recently reported the noncovalent DNA binding of syn-

[a] M. Pascu, Prof. M. J. Hannon
School of Chemistry, University of Birmingham
Edgbaston, Birmingham B15 2TT (UK)
Fax: (+44) 121-414-7871
E-mail: m.j.hannon@bham.ac.uk

[b] Dr. L. J. Childs, Dr. J. Malina, M. Pascu, M. J. Broome, Prof. P. M. Rodger, Prof. A. Rodger, Prof. M. J. Hannon
Department of Chemistry, University of Warwick
Coventry, CV4 7AL (UK)
Fax: (+44) 2476-524-112
E-mail: a.rodger@warwick.ac.uk

[c] Dr. M. J. Prieto, Prof. V. Moreno
Departament de Química Inorgànica
Diagonal 647, 08028-Barcelona (Spain)
Fax: (+34) 93-490-7725
E-mail: virtudes.moreno@qi.ub.es

[d] B. E. Rolfsnes, Prof. E. Sletten
Department of Chemistry, University of Bergen
Allgt. 41, 5007 Bergen (Norway)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author: Deconvolution of the film LD spectrum to assign Cu-helicate transition polarisations; calculation of the orientation of Cu-helicate on the DNA; DNA-cleavage activity of the cylinders with pBR322 plasmid DNA and hydrogen peroxide; DNA-cleavage activity of the cylinders with ct-DNA and hydrogen peroxide; possible major-groove and minor-groove binding orientations consistent with a 70° binding angle; NMR and ESI-MS data for the copper(I) cylinder.

thetic tetracationic metallosupramolecular cylinders that are similar in size and shape to protein DNA-recognition units.[5] These cylinders target the DNA major groove, spanning five or more base pairs, and induce dramatic intramolecular DNA coiling, which is unprecedented for synthetic agents and, in part, reminiscent of DNA coiling induced by histones in the cell nucleus.

DNA recognition by metal complexes is an area of intense investigation, driven primarily by the GG recognition, believed to be the key to the clinical activity of the metallodrugs cisplatin, carboplatin and oxaliplatin, which are among the most effective clinical anticancer agents.[6] This recognition takes place by the formation of platinum–nitrogen bonds.[7] However, metal complexes are also particularly suited for *noncovalent* approaches[7] and molecular designs, due to 1) the wide range of geometries and structures that metal units can support, 2) the ease of molecular assembly that coordination chemistry can afford, 3) the enhanced polarisation on coordination of hydrogen-bond acceptor and donor residues and, perhaps most importantly, 4) the cationic charge that the metallocentres impart to the reagents, which affords a substantial energetic contribution to the noncovalent binding to anionic DNA.

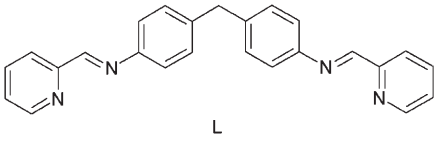
Examples of metal complexes that exhibit noncovalent interactions[3,8–12] with DNA include the octahedral ruthenium(II) complex $[Ru(phen)_3]^{2+}$ (phen = 1,10-phenanthroline) (although the precise binding modes of the Δ and Λ enantiomers of this compound seem complex)[9] and functionalised analogues that can act as metallointercalators, such as $[Ru(phen)_2(dppz)]^{2+}$ (dppz = dipyridophenazine). Barton has shown that the dppz ligand intercalates between base pairs of the DNA while the two phen ligands reside in a DNA groove,[10] and has recently applied related molecules to detect base-pair mismatches.[12] Nordén et al. reported bisintercalation for the complex $[Ru_2\{\mu\text{-}C4(cpdppz)_2\}(phen)_4]^{4+}$ (cpdppz = cyclopentadipyridophenazine). They suggest that the dppz groups of the molecule intercalate from either the major or minor groove with the bridging chain lying in the opposite groove.[11]

In previous studies[5] we used *tetracationic* cylinders[13] assembled by using metal dications. These cylinders are metallosupramolecular triple helicates,[14] in which three ligand strands are wrapped around two iron(II) centres. The molecular design is based on pyridylimine ligands and the cylinders can be assembled in one-pot reactions from commercial aldehydes, amines and metal salts. This allows the design to be varied systematically.[13,15,16] The various monometal complexes mentioned above are smaller than the dinuclear cylinders that we have developed and afford smaller molecular surfaces that span only two to three DNA base pairs. Moreover, none exhibit the dramatic intramolecular DNA coiling observed with the cylinder. Herein, we now describe the DNA binding of a dinuclear double-helical metallosupramolecular cylinder[16] assembled by copper(I) centres. We were intrigued to explore the effects of using copper(I) cylinders for two reasons. Firstly, the copper(I) cylinders would have a low charge, which might allow us to probe the effect of charge on strength of binding and extent of DNA coiling. Secondly, copper complexes of diimine ligands, such as phenanthroline, are known to exhibit oxidative DNA-cleavage activity that is thought to proceed by means of Fenton-generated hydroxyl-radical or copper-bound oxidants, such as $[CuO]^{+}$ or $[Cu(OH)]^{2+}$[17,18] (and there has been some recent interest[19] in polynuclear analogues). This opens up the exciting possibility that copper(I) cylinders might act as artificial nucleases.

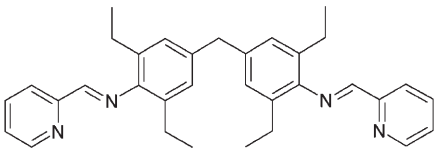
There has been a single report of DNA binding of copper(I) supramolecular double-helicates based on an oligobipyridine unit.[20] Binding to DNA was confirmed, although the precise binding mode was not unambiguously established. The dicationic copper(I) cylinders described herein bind strongly to DNA and do indeed cleave DNA in the presence of an oxidising agent. In contrast to mononuclear copper-containing artificial nucleases, such as the elegant Clip-Phen agents of Meunier,[18] which usually cleave a single DNA strand, these new copper(I) cylinders show a tendency to cleave both DNA strands at the same site.[21]

Results and Discussion

Our previous studies focused on the DNA binding of tetracationic cylinders with a triple-helical architecture formed by three ligand strands (structure **L**) wrapped around two metal dications.[5,13] Reaction of this ligand **L** with monocations, such as copper(I), leads to dinuclear double-stranded complexes with 2:2 stoichiometry $[M_2\mathbf{L}_2]^{2+}$.[15,16] However, in solution, these complexes are an equilibrium mixture of two dimeric isomers: a helicate (*rac*-isomer) and a metallocyclophane (*meso*-isomer) (Figure 1). To obtain exclusively the helical isomer, we introduced ethyl groups onto the spacer to give ligand $\mathbf{L}^{a}$.[16]



L



$\mathbf{L}^{a}$

Addition of ethyl groups to the central spacer destabilises the cyclophane configuration so that only $[M_2(\mathbf{L}^{a})_2]^{2+}$ helicate species are present in solution. This is evident from the ^{1}H NMR spectrum that reveals a single solution species at

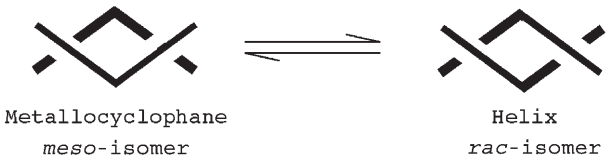


Figure 1. Representation of metallocyclophane (box) and helicate conformations.

room temperature and low temperature, and the central CH_2 resonance that being a singlet confirms the helical conformation (the cyclophane conformation gives rise to two doublets). Electrospray ionisation mass spectrometry (ESI-MS) confirms the dinuclear stoichiometry. The PF_6 salt has been crystallographically characterised and further confirms the double-helicate structure. This double-helical $[Cu_2(\mathbf{L}^a)_2]^{2+}$ cation represents a dicationic cylinder. Before detailing the DNA-binding studies it is pertinent to compare this copper cylinder with the previously studied iron(II), nickel(II) and cobalt(II) triple-stranded cylinders that bind in the major groove and induce the dramatic coiling (Figure 2).[5] These three triple-stranded cylinders are essentially isostructural and distances quoted are for the iron(II) cylinder only. The new copper(I) cylinder and the previously reported iron cylinder are of very similar length. Lengths along the cylinder axis between the extreme carbon atoms or the extreme hydrogen atoms[22] are: for copper 16.6 and 18.5 Å, respectively, and for iron 17.3 and 19.3 Å, respectively. Thus, the copper cylinder is very slightly shorter (~5%). The two cylinders are also similar in radius: for copper, radius-to-carbon is 4.5 Å and radius-to-hydrogen is 5.5 Å, and for iron, radius-to-carbon is 4.3 Å and radius-to-hydrogen is 5.3 Å. Thus, the copper cylinder is slightly wider (~4%). When viewed down the metal–metal axis (Figure 2b) the copper cylinder is more square. Thus, the copper cylinder has very similar dimensions to the iron(II) cylinder, but is slightly different in shape.

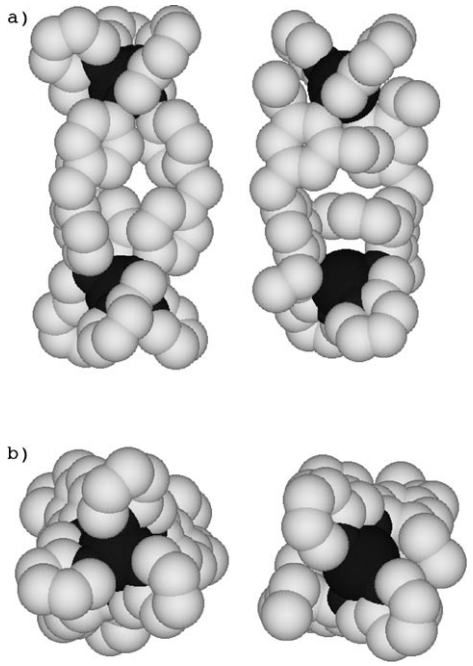


Figure 2. Space-filling side views (a) and end views (b) of the iron(II) cylinder[13] $[Fe_2(\mathbf{L})_3]^{4+}$ (left) and the copper(I) cylinder $[Cu_2(\mathbf{L}^a)_2]^{2+}$ (right); hydrogen atoms are omitted for clarity.

Absorption and circular dichroism spectroscopy: Circular dichroism (CD) was used as an initial screen to confirm the DNA binding of the copper cylinder. The CD and absorption spectra of calf-thymus DNA (ct-DNA) and $[Cu_2(\mathbf{L}^a)_2]^{2+}$ alone and together in solution in different mixing ratios were recorded. The UV-visible absorption spectrum of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ is shown in Figure 3 and reveals absorption bands with maxima at 320 and 480 nm; DNA has no absorbance above 300 nm. All of the spectra of the DNA/$[Cu_2(\mathbf{L}^a)_2]^{2+}$ solutions show cylinder absorption bands consistent with the increasing concentration of the cylinder, however, evidence of light scattering at higher cylinder loadings suggests some degree of DNA condensation or aggregation.

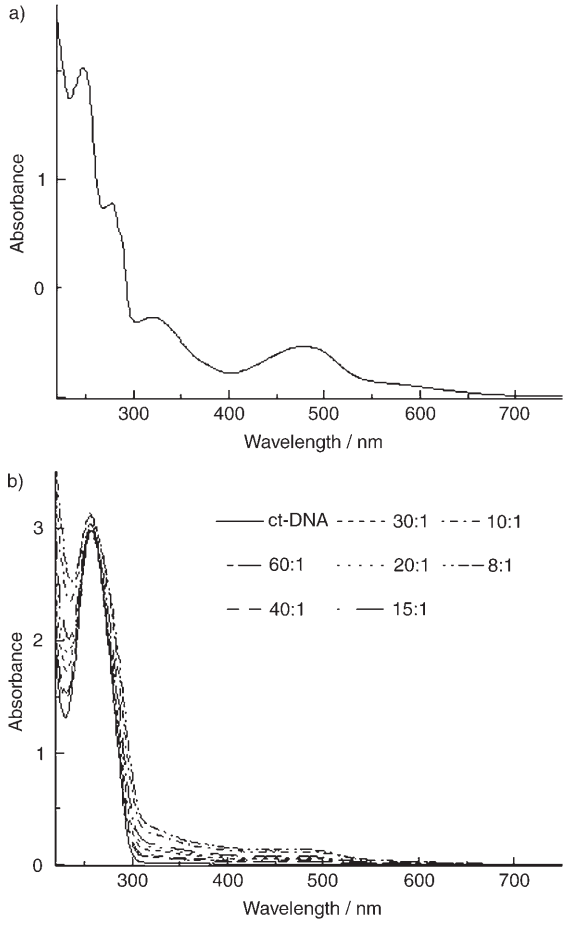


Figure 3. a) Absorption spectrum of 25 μM $[Cu_2(\mathbf{L}^a)_2]^{2+}$ in water. b) Absorption spectra of ct-DNA (500 μM DNA, 10 mM Na cacodylate, 20 mM NaCl) in the presence of $[Cu_2(\mathbf{L}^a)_2]^{2+}$. Mixing ratios ([DNA base]: $[Cu_2(\mathbf{L}^a)_2]^{2+}$]) are indicated. Pathlength = 1 cm.

As expected, $[Cu_2(\mathbf{L}^a)_2]^{2+}$ has no intrinsic CD signal because this compound exists as a racemic mixture of *P* and *M* helices. Therefore, any CD signal above 300 nm corresponding to cylinder absorbances that appears upon the addition of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ to DNA indicates interaction of the metal complex with the chiral DNA, as illustrated in Figure 4. The

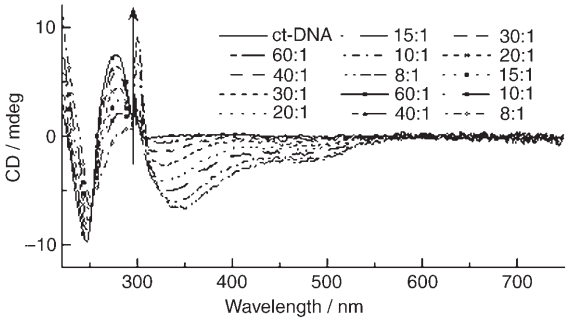


Figure 4. CD spectra of ct-DNA (500 μM DNA, 10 mM Na cacodylate, 20 mM NaCl) in the presence of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ in a 2-mm pathlength (295–220 nm) and in a 1-cm pathlength (750–295 nm). Mixing ratios ([DNA base]:$[Cu_2(\mathbf{L}^a)_2]^{2+}$]) are indicated.

weak, long-wavelength absorbance above 550 nm seen in the absorbance spectrum (Figure 3) is not apparent in the CD spectra. The induced CD (ICD) of the cylinder is negative from long wavelength to 305 nm, then shows a small positive signal from 295–305 nm, is negative from 252–295 nm and positive below this wavelength. The ICD in the DNA region could be due either to the presence of a competing ICD signal (from the cylinder) at the same wavelength, or to a ligand-induced DNA ICD, or to small structural changes in the DNA as a result of the helicate binding. Plots of ICD signal versus concentration of cylinder at 477, 345 and 270 nm (data not shown) all show smooth increases in signal, except for the last two data points in the DNA region, which is where there is evidence in all the spectra of DNA aggregation.

Flow linear dichroism: Linear dichroism (LD) is the difference between the absorption of linearly polarised light that is parallel to a chosen plane and that of linearly polarised light that is perpendicular to the chosen plane and can be used to probe the orientation of molecules. Long molecules, such as DNA (minimum length of ~250 base pairs), can be oriented in a flow Couette cell by viscous drag.[23] The linearly polarised light is incident radial to the flow cell and perpendicular to the flow direction. Small unbound molecules are not orientated in the experiment and show no signal. Similarly, molecules bound randomly to the DNA show no signal. However, molecules bound in a specific orientation with respect to the DNA will give a signal.

A flow LD titration series was carried out while keeping the DNA concentration constant at 500 μM (Figure 5). A negative signal was observed from 220–300 nm for the ct-DNA. This is characteristic of ct-DNA and is due to the base pairs lying approximately perpendicular to the DNA axis.[24] In the absence of DNA, the cylinder shows no signal,

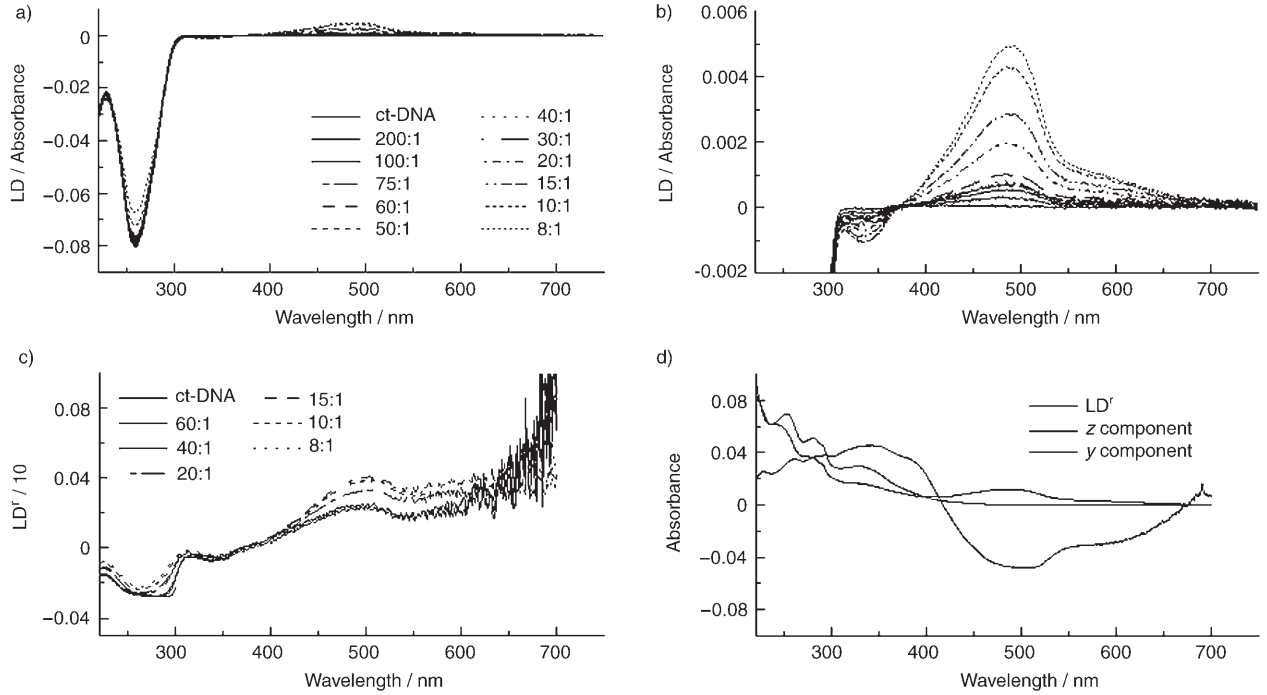


Figure 5. LD spectra of free ct-DNA (500 μM; 10 mM Na cacodylate, 20 mM NaCl) and in the presence of $[Cu_2(\mathbf{L}^a)_2]^{2+}$. Mixing ratios ([DNA base]: $[Cu_2(\mathbf{L}^a)_2]^{2+}$]) are indicated. a) Full scale, b) expanded scale. c) Reduced LD spectra (LD/Absorbance). Beyond 600 nm the absorbance tends to zero and LD^r becomes unreliable. d) Film LD^r spectrum of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ together with those of the *z* (orientation direction, long axis of cylinder) and *y* (perpendicular to *z*, i.e., in the *x*/*y* plane) component spectra.

being too small to be oriented by this flow method. A positive LD signal is observed from 400–700 nm (Figure 5b) upon the addition of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ to the DNA solution and a smaller negative signal is apparent for the in-ligand band at ~340 nm. The presence of these signals confirm that the cylinder is binding to the DNA in a specific orientation(s). As more $[Cu_2(\mathbf{L}^a)_2]^{2+}$ is added, the metal-to-ligand charge-transfer (MLCT) band increases, indicating that more helicate is binding to DNA upon each addition, in accord with the CD data. In contrast to the effect previously observed with the iron helicate, the copper cylinder has little bending effect on the DNA (little or no loss in the DNA signal at ~260 nm confirms DNA still orientated) until a ratio of ~10:1, at which scattering becomes significant in the absorbance and CD spectra (see above). We conclude, therefore, that the copper cylinder does not coil the DNA. The positive LD of the long-wavelength MLCT transition of the cylinder indicates that its transition moments are more parallel than perpendicular to the average DNA axis. By way of contrast, the in-ligand band at ~337 nm is negative and so polarised further from the DNA axis.

Because the DNA is not coiled by the copper cylinder, we can use the DNA LD to determine the orientation of the DNA and, hence, estimate the orientations of the cylinder transition moments on the DNA. The reduced LD (LD^r) is defined in Equation (1) in which α is the angle between the DNA helix axis and the transition moment polarisation and S is the orientation parameter.

$$LD^r = LD/Absorbance = 3S/2(3\cos^2\alpha - 1) \quad (1)$$

The value of α is approximately 86° for the DNA bases,[24] thus, $S \sim 0.18$ in this experiment. To analyse the LD data further, we need to understand something about the spectroscopy of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ itself. To study this, the copper cylinder was immobilised in a polyvinylalcohol (PVA) film that was then stretched. Under such conditions the long axis of the cylinder is preferentially aligned along the stretch direction, allowing component spectra to be determined from the LD data (Figure 5d). The method of deconvolution to component spectra is outlined in the Supporting Information. The long-wavelength MLCT band is almost completely polarised perpendicular to the helicate's long axis. This transition is polarised approximately along the line joining the Cu to the centre of the chelate. By contrast, the in-ligand region of ~340 nm is dominated by transitions moving along the ligands, which couple to give predominantly a transition polarised from the Cu to a point between the ligands. By taking the LD^r values from Figure 5c at the low ligand loading (no scattering; little contribution at 260 nm from cylinder spectroscopy), it follows that the long axis of the copper helicate lies at approximately 70° to the axis of the DNA helix (see Supporting Information for the calculation).

Fluorescence competition-binding assays: Both iron(II) and copper(I) cylinders bind strongly to cellulose,[25] which makes membrane dialysis unsuitable for the determination of binding constants. Therefore, to estimate the strength of binding, an ethidium bromide (EB) competition assay was carried out. This method monitors the displacement of EB from DNA by following the decrease in its fluorescence intensity as it is displaced from DNA into an aqueous environment. The concentrations of DNA and EB were kept constant throughout the experiment at a DNA:EB ratio of 4:5, and $[Cu_2(\mathbf{L}^a)_2]^{2+}$ was added sequentially. The cylinder displaces EB and the EB fluorescence is quenched. At a DNA base:cylinder ratio of about 15:1, the shape of the curve changes, presumably indicating the aggregation of the DNA. This occurs at a lower cylinder loading than in the CCD and LD experiments, because of the bound cationic EB. Thus, the copper cylinder does not completely displace the EB, in contrast to the situation with the iron cylinder. The data in Figure 6 allow us to conclude that the copper cylinder binds more strongly to DNA than does EB or $[Ru(phen)_3]^{2+}$ and less strongly than the iron cylinder. This confirms that (as would be expected) electrostatic charge is an important factor in determining the strength of binding to the anionic DNA (the tetracations bind much more strongly than the dication). Nevertheless, the shape is also a significant factor, with the cylindrical dication exhibiting a higher binding constant than the spherical $[Ru(phen)_3]^{2+}$, despite the fact that the latter will have a higher charge density. Thus, the shape (and the fit on the DNA) of the supramolecular cylinder is indeed important.

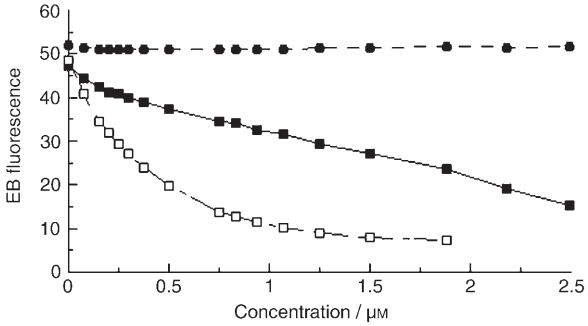


Figure 6. Fluorescence at 593 nm versus concentration of $[Cu_2(\mathbf{L}^a)_2]^{2+}$ (■), $[Fe_2(\mathbf{L})_3]$ (□) and $[Ru(phen)_3]^{2+}$ (●) in the presence of ct-DNA and EB (12 μM DNA, 15 μM EB, 10 mM Na cacodylate, 20 mM NaCl).

NMR studies: NMR titration experiments were performed to try to study the interaction between the copper(I) cylinder and a decamer $d(TATGGCCATA)_2$. Initial experiments involved the decamer (0.35 mM) dissolved in water with 10% of D_2O, 50 mM phosphate buffer and 40 mM NaCl. However, it became clear rapidly that the copper(I) cylinder was not stable in phosphate buffer (thus, phosphate buffer is not used in any of the other experiments described herein). To try to circumvent this problem, a second titration experiment was performed in which the decamer was dissolved in water with 10% D_2O and 30 mM NaCl at pH 5.95, at duplex concentration 0.4 mM. A 36 mM solution of the cylinder in 100% $[D_6]$DMSO was added to achieve a cylinder:duplex

ratio of 1:1. Although the cylinder was stable under these conditions, the addition led simply to precipitation of a red-brown solid that did not redissolve in water or aqueous methanol with ultrasound or heating. It seems likely that the cylinder precipitates the DNA at the concentrations required for NMR studies, which is consistent with the results at high loading ratios in the more-dilute CD, LD and UV-visible studies.

Artificial nuclease activity: To assess DNA-cleavage activity the interaction of the cylinder with pUC19 plasmid DNA was studied. Gel electrophoresis (1 % agarose gel) was used to visualise the effects (Figure 7). Incubation was for one

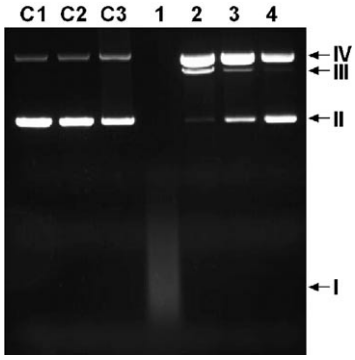


Figure 7. Gel electrophoresis of plasmid pUC19 after being incubated with $[Cu_2(\mathbf{L}^a)_2]^{2+}$ and hydrogenperoxide (HP). Lane C1, nonmodified plasmid; lane C2, plasmid with HP (hydrogen peroxide); lane C3, plasmid with $[Cu_2(\mathbf{L}^a)_2]^{2+}$ at 10:1 (base:cylinder) ratio; lanes 1—4, plasmid with HP and modified with $[Cu_2(\mathbf{L}^a)_2]^{2+}$ at different mixing ratios of 10:1, 20:1, 40:1 and 100:1, respectively. Band I, short fragments of DNA; band II, supercoiled plasmid; band III, linearised plasmid; band IV, relaxed plasmid. [HP] = 6.6 mM; [DNA] = 250 μM.

hour at 37 °C. The plasmid (lane C1) contains predominantly supercoiled DNA (band II) with small amounts of relaxed DNA (band IV). Addition of the cylinder to this plasmid DNA resulted in very little change in the gel chromatogram (lane C3). Similarly, addition of peroxide to the DNA caused no change (lane C2). However, addition of both cylinder and peroxide (lane 4) led to a reduction in supercoiled DNA (band II) and a corresponding increase in relaxed DNA (band IV). Relaxation of the supercoiling requires nicking of a single strand and, thus, implies that cylinder-induced DNA-strand scission occurs under these conditions. Increasing the concentration of cylinder (lanes 2, 3) leads to a further increase in relaxed DNA and a decrease in supercoiled DNA. Intriguingly, a band (band III) corresponding to linearised DNA (arising from cleavage of both strands at the same site) is apparent even in lane 4, well before the supercoiled DNA has disappeared. This is unusual (normally supercoiled DNA almost disappears before the formation of linear DNA) and could be explained as follows: 1) the complex has more than one reactive centre and cuts both DNA strands simultaneously; or 2) after the first cut, the complex stays bound and cuts the second strand; or 3) the binding of the copper cylinder is cooperative, leading to double-strand breaks. Although it is hard to be definitive, it seems that the ratios of the double-strand and single-strand breaks are independent of concentration, suggesting that the third proposal is unlikely. The first explanation is an attractive one, given the two-fold symmetry of the copper cylinder and the approximate two-fold symmetry of the DNA helix.

If the concentration of the cylinder is further increased (Figure 7, lane 1) the DNA is cleaved into short fragments (band I) that have high mobility within the gel. Control samples (free ligand with and without peroxide; and simple copper(II) chloride and hexafluorophosphate salts with and without peroxide) confirm that the complex is essential for cleavage. Very similar results were obtained with pBR322 plasmid (see Supporting Information).

A similar gel electrophoresis experiment with ct-DNA (a linear, polymeric DNA, as used in the spectroscopic experiments) confirmed that the effect is not merely restricted to circular DNAs. After addition of cylinder and peroxide to the ct-DNA, the bands corresponding to long DNA (closer to the loading well) disappeared, and only fast-running, short DNA fragments were observed (see Supporting Information).

Further confirmation of the DNA cleavage comes from linear dichroism experiments. As described, addition of the cylinder to ct-DNA has no effect on the DNA LD. Solutions incubated over periods of a few hours show no loss in DNA LD signal. However, if peroxide is added to the cylinder-DNA solution there is a fairly rapid loss of the DNA LD signal, consistent with cleavage of the ct-DNA into small fragments that can no longer be oriented in the flow cell. Peroxide alone has no effect on the ct-DNA LD signal.

Visual conformation of the DNA-cleavage activity comes from molecular-level images of the pBR322 plasmid DNA and copper(I) cylinder, obtained by using tapping-mode atomic force microscopy (AFM) (Figure 8). Plasmid samples incubated for one hour with cylinder alone show no evidence of strand scission (Figure 8a, b), whereas the sample incubated with 6 μM copper(I) cylinder and hydrogen peroxide shows clear evidence of DNA-strand scission to give linear fragments (Figure 8c). This effect is even more dramatic upon addition of a higher concentration of copper(I) cylinder (20 μM) to the DNA (Figure 8d). In this case, we observe just short pieces of DNA as a result of the cleavage activity. Extended incubation times led to small fragments even at lower cylinder concentration.

Conclusion

The results obtained with this dicationic copper(I) cylinder cast further light on the role of charge in the binding and coiling of DNA by tetracationic cylinders.

Charge is important in binding (as expected), and tetracationic cylinders bind more strongly than this dicationic cylinder. Indeed, electrostatic forces will be important in designing noncovalent DNA-recognition agents that bind

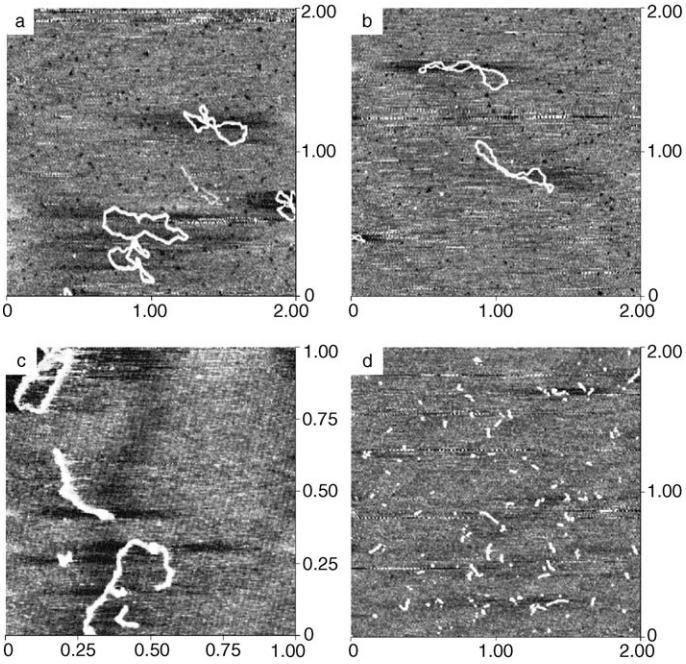


Figure 8. pBR322 DNA with (a) 6 μM copper(I) cylinder, (b) 20 μM copper(I) cylinder, (c) 6 μM copper(I) cylinder with H_2O_2 and (d) 20 μM copper(I) cylinder with H_2O_2. All samples were incubated for 1 hr at 20 °C. The dimensions are in μm.

strongly to DNA. However, the molecular shape is also important, with the cylindrical dication exhibiting stronger binding than spherical $[Ru(phen)_3]^{2+}$, despite a lower charge density. This validates our original hypothesis[5] that benefits could be accrued by scaling up from small molecules to designing supramolecular arrays with size and shape similar to nature's own DNA-recognition motifs. The results also imply that charge may be important in the DNA bending and coiling observed with the tetracationic iron(II) cylinder. Our dicationic copper(I) cylinder does not cause such coiling. Charge has been proposed as an important factor in DNA-bending processes[26] and indeed is believed to play a role in DNA coiling about histones. The copper cylinder induces DNA aggregation (leading to precipitation) at mixing ratios much lower than those for the tetracationic cylinders, despite its lower binding constant, and this may be related to the lack of observed intramolecular DNA coiling (in the intramolecular coils the cylinders are wrapped up by the DNA, so they cannot bridge between and aggregate DNA molecules).

The copper cylinder binds to DNA without disturbing its B-DNA configuration. Based on the size of the cylinder and the pitch of the DNA, a cylinder lying perfectly along the major groove would make an angle between its long axis and the axis of the DNA helix of around 60°. The observed average binding angle (~70°) of this copper(I) cylinder would, therefore, be consistent with binding in the major groove or with a minor-groove binding mode (which, due to the cylinder size, might involve lying outside the minor groove, perhaps with partial insertion of the cylinder into the groove). Support for this comes from a preliminary molecular-dynamics simulation[27] of poly-d(AT)$_2$ focused on configurations in which the cylinder made angles of 70° (±5°) with the DNA axis. This revealed several binding modes in which the cylinder sat nicely in the major groove, together with one configuration outside the minor groove. These potential binding modes are visualised in images provided in the Supporting Information, although these are not intended to imply definitive binding modes.

Interestingly, this copper cylinder exhibits DNA-cleavage activity in the presence of peroxide. The cylinder exhibits an unusual tendency to perform a double-strand cleavage at the same site, possibly reflecting its dinuclear nature. The cleavage ability extends the potential applications of these metallosupramolecular cylinders, opening up the possibility of using copper-based cylinders as artificial nucleases. Studies are currently underway to impart sequence selectivity to the cylinder systems for this purpose.

Experimental Section

Materials: Ultrapure water (18.2 MΩ) was used in all experiments. The ct-DNA (highly polymerised) was purchased from Sigma-Aldrich and was dissolved in water without any further purification. Stock solutions of ct-DNA were kept frozen until the day of use. DNA solutions were prepared as required from the frozen samples by diluting to the desired concentration of polynucleotide with a buffer of 10 mM cacodylate (prepared from sodium cacodylate adjusted to pH 6.8 with hydrochloric acid) and 20 mM NaCl. The DNA concentrations were determined spectroscopically by using the known molar-extinction coefficient of $\varepsilon_{258} = 6600\ mol^{-1}\,dm^3\,cm^{-1}$ per DNA base.[28] Commercially available tris-acetate-EDTA (TAE, from Fisher) working buffer was used for gel electrophoresis of pUC19 plasmid DNA (New England Biolabs). The gel loading buffer was prepared by dissolving 0.25% (w/v) of xylene cyanole, 0.25% (w/v) of bromophenol blue and 30% (w/v) of glycerol in water.

Preparation of ligand L^a: This was prepared by mixing two equivalents of pyridine-2-carboxyaldehyde and one equivalent of 4,4′-methylene-bis(2,6-diethylaniline) in methanol as previously reported.[16] Preparation of the chloride salt of the copper(I) complex was analogous to that of the corresponding PF_6 salt, which has been characterised crystallographically:[16] $[Cu_2(\mathbf{L}^a)_2][Cl]_2$ was prepared by mixing ligand **L^a** and copper(I) chloride in a 1:1 ratio in methanol. The dark-red solution was heated under reflux overnight and then cooled to RT. The solvent was removed by rotary evaporation and the red solid was dried over P_2O_5. ^{1}H NMR (MeOH, 298 K): $\delta = 8.79$ (s, 1H; H_i), 8.62 (d, $J = 4.5$ Hz, 1H; H_6), 8.28, (t, $J = 7.3$ Hz, 1H; $H_{4/5}$), 8.14 (d, $J = 7.5$ Hz, 1H; H_3), 7.87 (br dd, $J = 7.0$, 5.3 Hz, 1H; $H_{4/5}$), 7.07 (br s, 1H; H_{Ph}), 6.57 (br s, 1H; H_{Ph}), 3.87 (s, 1H; central CH_2), 2.73 (br m, 1H; CH_2), 2.64 (br m, 1H; CH_2), 2.21 (br m, 1H; CH_2), 2.04 (br m, 1H; CH_2), 1.04 (br s, 3H; CH_3), 0.56 ppm (br s, 3H; CH_3); UV/Vis (H_2O): λ_{max} (ε) = 258 (6600), 478 (9450), 321 (14800), 278 nm (36800 $mol^{-1}\,dm^3\,cm^{-1}$); ESI-MS (MeOH): m/z: 552 $[Cu_2(\mathbf{L}^a)_2]^{2+}$.

Circular dichroism: Spectra were collected in cuvettes of pathlength 2 mm and 1 cm by using a Jasco J-715 spectropolarimeter. Spectroscopic titrations were performed from which CD and UV/Vis absorption spectra were recorded. Titrations were carried out at constant concentrations of DNA (500 μM), NaCl (20 mM) and sodium cacodylate buffer (10 mM). The ratio of DNA:metal complex was decreased during the titration series by incrementing the concentration of metal complex in the cuvette from 0–62.5 μM. Two stock solutions were prepared. The first was of metal complex in water (500 μM) and the second solution contained DNA (1000 μM), NaCl (40 mM) and sodium cacodylate buffer (20 mM). After the addition of x cm^3 of the metal-complex solution to the cuvette, an equivalent volume of the second DNA stock solution was added. This

meant that the concentrations of DNA, NaCl and sodium cacodylate buffer in the cuvette remained unaltered.[29]

Linear dichroism: Flow LD spectra were collected by using a flow Couette cell in a Jasco J-715 spectropolarimeter adapted for LD measurements. Long molecules, such as DNA (minimum length of ~250 base pairs), can be orientated in a flow Couette cell. The flow cell consists of a fixed outer cylinder and a rotating solid quartz inner cylinder, separated by a gap of 0.5 mm, giving a total pathlength of 1 mm.[23,29]

Stretch film is used to orient small molecules and involves dissolving the molecule of interest in a polymer solution, such as polyvinylalcohol (PVA). A 10% (w/v) low-molecular-weight PVA solution in water (4.8 cm^3) was prepared and heated to near boiling to ensure that all the PVA had dissolved. The solution was cooled and a saturated aqueous solution of metal complex (0.2 cm^3) added. The viscous solution was then cast onto a glass plate and left to dry in a ventilated, dust-free environment. A blank film was also prepared by adding water (0.2 cm^3) instead of metal complex to the PVA solution. When dry, both films were carefully peeled from the glass plate by using a scalpel and were then placed in a mechanical stretcher. Under gentle heating the films were stretched by a factor of two. The UV/Vis and LD spectra of both films were recorded and the spectra of the blank film was subtracted from that of the metal-complex film. It was then possible to calculate component spectra (see Supporting Information).

Fluorescence competition assay: An ethidium bromide (EB) competition-binding assay was carried out according to a modified literature method[30] by using a Perkin–Elmer LS-50B spectrofluorimeter.

Gel electrophoresis: All electrophoresis experiments were carried out by using a Pharmacia GNA-100 submarine unit. Gel trays of 110×100 mm were used, with an 11-toothed comb to produce the sample wells. A Pharmacia Electrophoresis Power Supply, ECPS-3000, was used as a constant voltage supply set to 70 V and 65 mA. The volume of agarose solution to be used was calculated so as to produce a gel with a depth of 4 mm. TAE working buffer was used for the gel electrophoresis. Gels were stained with ethidium bromide visualised under a UV lamp and were photographed by using a UVP white/UV transilluminator. The samples were prepared by mixing the metal complex with DNA in different ratios, keeping the DNA concentration constant at 2.5×10^{-4} M. For the DNA cleavage studies, all samples were incubated for 1 hr at 37°C in 1 mM sodium cacodylate and 20 mM NaCl prior to analysis by gel electrophoresis.

Atomic force microscopy:[31] AFM images of pBR322 plasmid DNA were collected by using two different concentrations of copper(I) cylinder (6 and 20 μM). DNA–metal-complex adducts were prepared as follows: pBR322 DNA was incubated in an appropriate volume with the required metal complex in 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES) buffer (HEPES 4 mM pH 7.4, KCl 5 mM, $MgCl_2$ 2 mM). All solutions were prepared with water (18.2 MΩ) that had been filtered through 0.2 nm FP030/3 filters (Schleider and Schuell, Germany) and centrifuged at 4000 g several times to avoid salt deposits and to provide a clear background upon imaging. The samples were left to equilibrate at 37°C for 30 min in the dark. Samples were prepared for AFM by placing a drop (6 μL) of DNA–metal-complex adduct solution onto mica (Ashville–Schoonmaker Mica, Newport News, VA). After adsorption for 5 min at RT, the samples were rinsed for 10 s in a jet of deionised water (18 $M\Omega\,cm^{-1}$ Milli-Q water) directed from a squeeze bottle onto the surface. The samples were blow-dried with compressed argon over silica gel and then imaged by using a Nanoscope III Multimode AFM (Digital Instrumentals, Santa Barbara, CA) operating in tapping mode in air at a scan rate of 1–3 Hz. The AFM probes were 125 μm long monocrystalline silicon cantilevers with integrated conical-shaped Si tips.

Acknowledgements

We thank the Leverhulme Trust (Grant F/125/BC) and the EU (MARCY RTN HPRN-CT-2002-00175; Marie Curie fellowship HPMF-CT-2002-01856) for support. This work was conducted within the framework of COST D20 (metal complexes in anticancer and antiviral therapy), working group D2O/0010/02 (noncovalent DNA recognition), which supported an exchange of researchers (B.E.R.). We thank the EPSRC National Mass Spectrometry Service Centre, Swansea, for recording electrospray mass spectra. M.J.H. is the Royal Society of Chemistry Sir Edward Frankland Fellow 2004–2005.

[1] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, Freeman, New York, **2002**; S. Neidle, *Nucleic Acid Structure and Function*, Oxford University Press, Oxford, **2002**.

[2] C. Brandon, J. Tooze, *Introduction to Protein Structure*, Garland, New York, **1998**.

[3] P. B. Glover, P. R. Ashton, L. J. Childs, A. Rodger, M. Kercher, R. M. Williams, L. De Cola, Z. Pikramenou, *J. Am. Chem. Soc.* **2003**, *125*, 9918–9919; corrigendum: P. K. L. Fu, P. M. Bradley, C. Turro, *Inorg. Chem.* **2003**, *42*, 878–884; corrigendum: *Inorg. Chem.* **2004**, *43*, 2220; C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton, D. C. Rees, *Nat. Struct. Biol.* **2000**, *7*, 117–121; I. Greguric, J. R. Aldrich-Wright, J. G. Grant Collins, *J. Am. Chem. Soc.* **1997**, *119*, 3621–3622; J. L. Kisko, J. K. Barton, *Inorg. Chem.* **2000**, *39*, 4942–4949; G. Bobba, J. C. Frias, D. Parker, *Chem. Commun.* **2002**, *8*, 890–891; G. Colmenarejo, A. Holmen, B. Nordén, *J. Phys. Chem. B* **1997**, *101*, 5196–5204.

[4] C. Melander, D. M. Herman, P. B. Dervan, *Chem. Eur. J.* **2000**, *6*, 4487–4497; P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215.

[5] M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 880–884; I. Meistermann, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, S. Khalid, P. M. Rodger, J. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074; E. Molderheim, M. J. Hannon, I. Meistermann, A. Rodger, E. Sletten, *J. Biol. Inorg. Chem.* **2002**, *7*, 770–780; M. J. Hannon, A. Rodger, *Pharmaceutical Visions*, **2002**, *Autumn Issue*, 14–16; A. Oleski, A. G. Blanco, R. Boer, I. Usón, J. Aymami, A. Rodger, M. J. Hannon, Miquel Coll, *Angew. Chem.* **2006**, *118*, 1249–1253; *Angew. Chem. Int. Ed.* **2006**, *45*, 1227–1231.

[6] B. Lippert, *Cisplatin, Chemistry and Biochemistry of A Leading Anticancer Drug*, Wiley-VCH, Weinheim, **1999**; J. Reedijk, *Chem. Commun.* **1996**, 801–806; Z. J. Guo, P. J. Sadler, *Adv. Inorg. Chem.* **2000**, *49*, 183–306; J. D. Roberts, J. Peroutka, N. Farrell, *J. Inorg. Biochem.* **1999**, *77*, 51–57; S, J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA, **1994**.

[7] Cisplatin binds to the DNA bases through coordination bonds (not covalent bonds), yet the binding is in some ways analogous to that of DNA alkylating agents, such as nitrogen mustards, with which they are classified clinically. With the term "noncovalent" binding by metal complexes used herein we are formally referring to noncovalent binding that does not involve direct coordination of the metal centre to atoms on the DNA molecule.

[8] K. K. Patel, E. A. Plummer, M. Darwish, A. Rodger, M. J. Hannon, *J. Inorg. Biochem.* **2002**, *91*, 220–229; J. Aldrich-Wright, C. Brodie, E. C. Glazer, N. W. Luedtke, L. Elson-Schwab, Y. Tor, *Chem. Commun.* **2004**, 1018–1019; S. E. Wellman, D. B. Sittman, J. B. Chaires, *Biochemistry* **1994**, *33*, 384–388.

[9] C. V. Kumar, J. K. Barton, N. J. Turro, *J. Am. Chem. Soc.* **1985**, *107*, 5518–5523; C. Hiort, B. Nordén, A. Rodger, *J. Am. Chem. Soc.* **1990**, *112*, 1971–1982; M. Eriksson, M. Leijon, C. Hiort, B. Nordén, A. Graslund, *J. Am. Chem. Soc.* **1992**, *114*, 4933–4934.

[10] K. E. Erkkila, D. T. Odom, J. K. Barton, *Chem. Rev.* **1999**, *99*, 2777–2795.

[11] B. Önfelt, P. Lincoln, B. Nordén, *J. Am. Chem. Soc.* **2001**, *123*, 3630–3637.

[12] U. Schatzschneider, J. K. Barton, *J. Am. Chem. Soc.* **2004**, *126*, 8630–8631; E. Ruba, J. R. Hart, J. K. Barton, *Inorg. Chem.* **2004**, *43*, 4570–4578.

[13] M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin, W. Errington, *Chem. Commun.* **1997**, 1807–1808.
[14] J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCH, Weinheim, **1995**; C. Piguet, G. Bernardinelli, G. Hopfgartner, *Chem. Rev.* **1997**, *97*, 2005–2062; M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3497; M. J. Hannon, L. J. Childs, *Supramol. Chem.* **2004**, *16*, 7–22.
[15] M. J. Hannon, C. L. Painting, N. W. Alcock, *Chem. Commun.* **1999**, 2023–2024; L. J. Childs, N. W. Alcock, M. J. Hannon, *Angew. Chem.* **2001**, *113*, 1113–1115; *Angew. Chem. Int. Ed.* **2001**, *40*, 1079–1080; L. J. Childs, N. W. Alcock, M. J. Hannon, *Angew. Chem.* **2002**, *114*, 4418–4421; *Angew. Chem. Int. Ed.* **2002**, *41*, 4244–4247.
[16] L. J. Childs, M. Pascu, A. J. Clarke, N. W. Alcock, M. J. Hannon, *Chem. Eur. J.* **2004**, *10*, 4291–4300.
[17] D. S. Sigman, T. W. Bruice, A. Mazumder, C. L. Sutton, *Acc. Chem. Res.* **1993**, *26*, 98–104.
[18] a) M. Pitie, B. Donnadieu, B. Meunier, *Inorg. Chem.* **1998**, *37*, 3486–3489; b) M. Pitie, C. Boldron, H. Gornitzka, C. Hemmert, B. Donnadieu, B. Meunier, *Eur. J. Inorg. Chem.* **2003**, 528–540.
[19] S. T. Frey, H. H. J. Sun, N. N. Murthy, K. D. Karlin, *Inorg. Chim. Acta* **1996**, *242*, 329–338; K. J. Humphreys, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.* **2002**, *124*, 6009–6019; K. J. Humphreys, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.*, **2002**, *124*, 8055–8066; C. Tu, Y. Shao, N. Gan, D. Xu, Z. J. Guo, *Inorg. Chem.* **2004**, *43*, 4761–4766.
[20] B. Schoentjes, J.-M. Lehn, *Helv. Chim. Acta* **1995**, *78*, 1–12.
[21] Although mononuclear complexes usually favour single-strand cleavage, double-strand cleavage is sometimes also possible: See, for example, ref. [18b] and P. U. Maheswari, S. Roy, H. den Dulk, S. Barends, G. van Wezel, B. Kozlevcar, P. Gamez, J. Reedijk, *J. Am. Chem. Soc.* **2006**, *128*, 710–711.
[22] Lengths quoted are the distances between the centroids of the three pyridine C5 atoms at each end of the helix or the three pyridine H5 atoms. Radii are taken perpendicular to the helix axis.
[23] A. Rodger, B. Norden, *Circular and Linear Dichroism*, Oxford University Press, Oxford, **1997**.
[24] P.-J. Chou, W. C. Johnson, *J. Am. Chem. Soc.* **1993**, *115*, 1205–1214.
[25] M. J. Hannon, I. Meistermann, C. J. Isaac, C. Blomme, J. Aldrich-Wright, A. Rodger, *Chem. Commun.* **2001**, 1078–1079.
[26] L. J. Maher, III, *Curr. Opinion Chem. Biol.* **1998**, *2*, 688–694; I. Rouzina, A. Bloomfield, *Biophys. J.* **1998**, *84*, 3152–3564.
[27] These initial simulations are used merely to visualise how a 70° binding angle might appear. Detailed molecular-dynamics simulations and their analysis are beyond the scope of this current study, but will be described in a future report. For details of a related molecular-dynamics simulation of the iron(II) cylinder on DNA see S. Khalid, A. Rodger, M. J. Hannon, P. M. Rodger, *Chem. Eur. J.* **2006**, *12*, 3493–3506
[28] R. D. Wells, J. E. Larson, R. C. Grant, B. E. Shortle, C. R. Cantor, *J. Mol. Biol.* **1970**, *54*, 465–497.
[29] A. Rodger, *Methods Enzymol.* **1993**, *226*, 232–258.
[30] A. McCoubrey, H. C. Latham, P. R. Cook, A. Rodger, G. Lowe, *FEBS Lett.* **1996**, *380*, 73–78.
[31] G. B. Onoa, G. Cervantes, V. Moreno, M. J. Prieto, *Nucleic Acids Res.* **1998**, *26*, 1473–1480.

Received: January 16, 2006
Published online: April 27, 2006

WILEY InterScience® DISCOVER SOMETHING GREAT CHIRALITY 18:680–690 (2006)

A New Method for Fibrous Protein Analysis Illustrated by Application to Tubulin Microtubule Polymerisation and Depolymerisation

RACHEL MARRINGTON,[1] MARK SEYMOUR,[2] AND ALISON RODGER[1]*
[1]*Department of Chemistry, University of Warwick, Coventry, United Kingdom*
[2]*Syngenta, Jealott's Hill, Bracknell, Berkshire, United Kingdom*

Presented at the 10th International Conference on Circular Dichroism, 2005, Sandestin, Florida

ABSTRACT A thermostatted micro volume Couette cell has been designed to enable linear dichroism (*LD*) data to be collected at a range of temperatures. The cell is a development of the traditional Couette flow *LD* cell and includes the recent development of micro-volume *LD* (20–40 μL) coupled with the addition of a heating element, temperature probe and controller. This new micro volume Couette *LD* cell opens the way not only to the *LD* analysis of systems where sample volume is critical, but also for the *LD* analysis of temperature sensitive samples. The polymerization of the microtubule protein tubulin has been followed in a range of different conditions using the thermostatted micro volume Couette *LD* cell. The focusing lenses on the cell, which are required for the microvolume cell, have the side benefit of significantly reducing the light-scattering artifacts caused by the large size of tubulin microtubules. It is now possible to monitor real-time polymerization and depolymerization kinetics, and any structural rearrangements of chromophores within the polymer. In the case of tubulin, the *LD* spectra revealed a greater change in the orientation of tryptophan residues at ~290 nm during polymerization compared to other contributing chromophores—guanine, phenylalanine, and tyrosine. The improvements in instrumental design have also allowed *LD* spectra of tubulin to be collected down to ~230 nm (previous data have only been available from the near UV region), which means that some indication of protein backbone-orientation changes are now available. It was observed during this work that apparent *LD* intensity maxima are in fact artifacts when the high-tension voltage is high. The onset of such artifacts has been observed at much lower voltages with light-scattering fibrous proteins (including tubulin) than with nonscattering samples. Therefore, caution must be used when interpreting *LD* data collected with medium to high photomultiplier tube voltages. *Chirality 18:680–690, 2006.*

KEY WORDS: thermostatted; microvolume; linear dichroism; Couette; tubulin; kinetics; polymerization; depolymerization

Microtubules, a key structural element of eukaryotic cells, are ~25-nm-wide hollow robust helical polymers composed of the heterodimeric protein tubulin.[1,2] They have a range of structural and dynamic roles in the cell. Microtubules are directly involved in mitosis (cell division) and organization of components within the cell, for example, the movement of organelles. Microtubules are highly dynamic and the control of polymerization is central to their biological function, spatial arrangements, and response to cellular needs.[2] Studies of tubulin have been carried out *in vitro* and have shown that pure tubulin of a high enough concentration will polymerize into microtubules at 37°C as long as the cofactors Mg^{2+} and GTP (guanosine triphosphate) are present.[3] Initially, tubulin heterodimers nucleate into oligomers which then form protofilaments. The microtubules grow by an elongation process involving the addition of tubulin heterodimers. A single protofilament is thermally unstable; however, lateral interactions allow the protofilaments to line up side by side into a sheet of protofilaments which then curl around to form a hollow tube. A process called 'treadmilling' occurs when tubulin heterodimers add to one end of the growing microtubule and tubulin heterodimers are removed from the opposite end. The displaced tubulin heterodimer is then able to repeat the cycle until all the GTP in the system has been utilized. Depolymerization occurs once this reservoir of GTP has been depleted or at temperatures below 37°C. Traditionally, light scattering has been used to monitor the polymerization of tubulin into microtubules.[4] A change in the UV 'absorbance' at, for

*Correspondence to: Alison Rodger, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, United Kingdom.
E-mail: a.rodger@warwick.ac.uk
Received for publication 2 December 2005; Accepted 14 April 2006
DOI: 10.1002/chir.20305
Published online 5 July 2006 in Wiley InterScience (www.interscience.wiley.com).

example, 350 nm (a wavelength where microtubule proteins have no true absorption, but where scattering prevents photons from reaching the detector, thus giving an apparent absorbance signal) is due to light scattering by the growing microtubules. There are two examples in the literature where linear dichroism (*LD*) has been used to study tubulin microtubules.[5,6] In this work we show that not only can *LD* be used instead of absorbance, but it provides additional information about structural changes during polymerization and microtubule formation.

Various drugs have been developed to affect microtubule assembly, commonly called antimitotic drugs. Drugs which interact with tubulin are useful tools for cell biologists to probe the role and regulation of microtubules in the cell. For a detailed review, see Jordan and Wilson.[7] Taxol™, also known as paclitaxel, is an antitumour drug extracted from the bark of yew trees. In the early 1980s Taxol™ was found to bind reversibly to microtubules (rather than the tubulin heterodimer) and promote the nucleation and elongation phases of polymerization.[8] It has been used as a drug in the treatment of some cancers, particularly breast cancer, since the early 1990s, as it promotes the formation of microtubules, thus blocking the division of cancer cells.

Probably the most well-known drug that works by inhibition of tubulin polymerization is colchicine. This is a plant alkaloid that has been used medicinally in the treatment of gout. Colchicine binds tightly and reversibly in a 1:1 ratio to the tubulin heterodimer (i.e., not the microtubule). When a colchicine-bound tubulin heterodimer adds to the growing end of the microtubule, it blocks further polymerization, thus (because of the 'treadmilling' process) resulting in microtubule disassembly.[9] Vincristine is another antimitotic antitumour drug. It has been reported that vincristine (and other members of the vinca alkaloid family) aggregate tubulin *in vitro*, and can bind to microtubules both along its walls and at both ends. This results in the microtubules protofilaments separating, stabilizing, and subsequently coiling.[10–12]

Understanding how all these molecules directly interact with microtubules and the structural implications of their binding could contribute significantly to the design and evaluation of new molecules for cell biology and neuroscience. This input requires new methods to study the structure and kinetics of tubulin. The thermostatted micro volume Couette *LD* cell we have developed is ideally suited for this purpose.

LD is defined to be the difference in absorbance of light linearly polarized parallel and perpendicular to an orientation axis, given by eq. (1) below. If a transition moment is aligned more parallel than perpendicular to the orientation axis, a net positive *LD* signal is observed. When a transition moment is at an angle of 54.7° the *LD* signal equals zero, however effective the orientation.

$$LD = A_{\parallel} - A_{\perp} \qquad (1)$$

Although one can measure *LD* by inserting polarizers into the light beam of a normal absorption spectrometer, the more expensive option of using a circular dichroism spectrometer is attractive due to significantly improved sensitivity.[13] Circularly polarized light can be converted to a linearly polarized light by the addition of a quarter-wave plate so *CD* spectropolarimeters can readily be converted to *LD* instruments. Alternatively, one may increase the voltage on the photoelastic modulator and directly produce the required alternating linearly polarized light beams. While *CD* requires the use of chiral compounds to generate a net *CD* signal, *LD* requires systems which are either intrinsically oriented or can be oriented during an experiment.

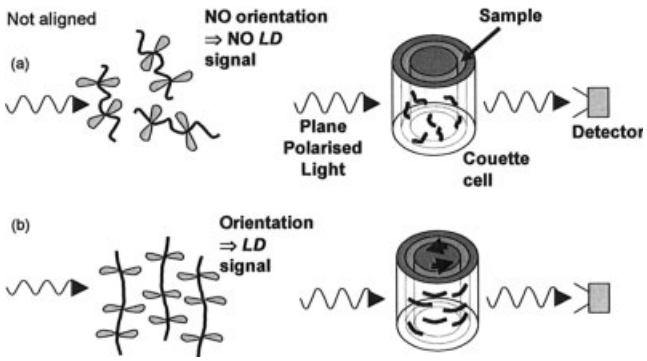


Fig. 1. Schematic diagram of an *LD* Couette cell: (a) stationary inner and outer cylinders therefore no sample orientation and no net *LD* signal; (b) rotating inner and stationary outer cylinder resulting in sample orientation and thus an *LD* signal.

Quantitative information about the orientation of a molecule can be obtained from the reduced linear dichroism LD^r, given by eq. (2) below. A is the absorption of the sample under anisotropic conditions (i.e., not oriented) and S is a scaling factor (the orientation factor) that defines the efficiency of the macroscopic orientation. S would equal 1 for perfect orientation and 0 for random orientation. α is the angle that the transition moment responsible for the absorption of light at a particular wavelength makes with the orientation axis. If either S or α is known, the other can be calculated.

$$LD^r = \frac{LD}{A} = \frac{A_{\parallel} - A_{\perp}}{A} = \frac{3}{2}\mathrm{S}(3\cos^2\alpha - 1) \qquad (2)$$

LD requires a method of orientation, the most common being Couette flow (Fig. 1), where the solution is held between two concentric cylinders and the molecules are aligned by the viscous drag that occurs when one of the cylinders (outer or inner) rotates and the other remains stationary. (The differences between inner and outer rotating systems has been discussed previously.[14,15]) The Couette cell employed in this work (Fig. 2) utilizes an outer rotating cylinder (quartz capillary) and a stationary inner cylinder (quartz rod). The light must only be incident on the middle front and back of the solution, otherwise an averaging over molecules parallel and perpendicular to the propagation direction would occur. So for this capillary cell, which is designed to fit into a large compartment Jasco CD machine whose light beam is 8–10 mm in diameter, focusing lenses are incorporated before and after the sample in order to maximize the photon count incident on the sample and also on the photomultiplier tube. The post-sample lens has the

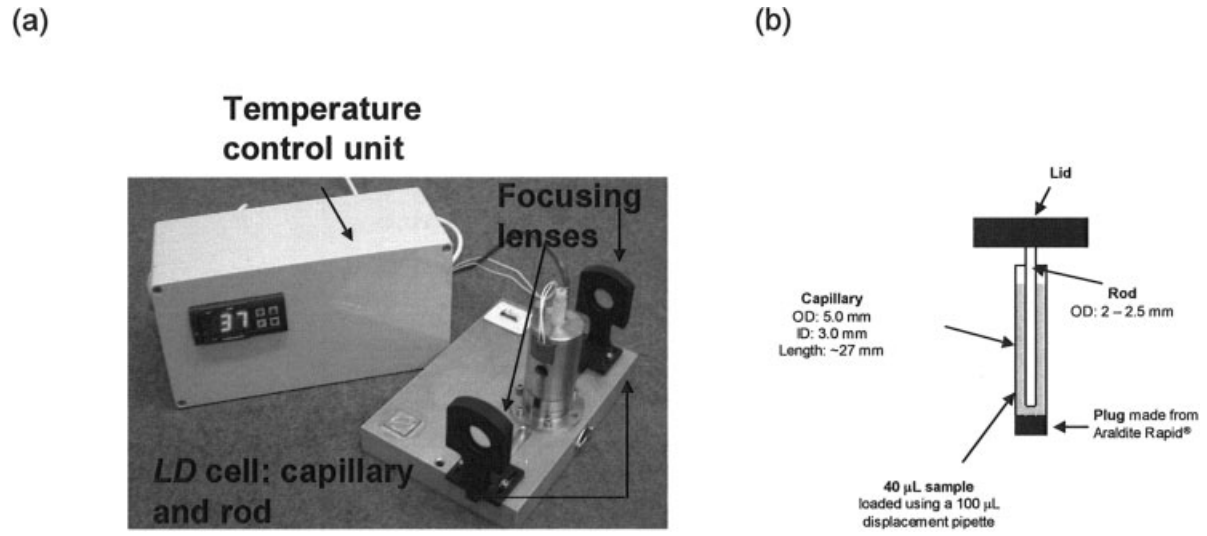


Fig. 2. (a) Photograph of thermostatted micro volume Couette *LD* cell; (b) schematic diagram showing the capillary and rod assembly in the micro volume Couette *LD* cell.

added advantage of significantly reducing the scattering of unabsorbed light by the sample. The microtubules formed still give rise to some light scattering but, as described below, this can be corrected by using a method described by Nordh et al.[5] (see Materials and Methods section).

Couette flow *LD* was initially developed in the early 1960s by A. Wada.[16] Its main application to date has been for DNA analysis; however, in the 1960s–1970s, flow *LD* was recognized as a key technique in fibrous protein analysis.[17–20] Limitations in both the Couette cells available at that time and also spectrometer signal to noise restricted the data collection on fibrous proteins to the near UV wavelength range. Our recent developments in Couette flow *LD* have reduced the sample volume from ~2–3 mL in the traditional Couette *LD* flow systems, to ~20 μL and enabled data collection down into the far UV region of the spectrum.[21,22] Information about the design and validation of the unthermostatted analogues of the cell used in this work system has been described previously.[15,22,23]

The work reported here has required a further development of the micro volume Couette *LD* cell in that it is now possible to thermostatically control the temperature of the solution within the capillary to within ~0.2% of the set °C temperature by the addition of a heating element, temperature probe, and controller. The capillary unit is encased within a thermostatted metal jacket with silicon rubber and fiberglass heater pads. Temperatures between room temperature and ~60°C can be achieved. It is important to note that the thermal energy generated by the motor (which is located directly below the capillary in this design) needs to be taken into account. The increase in sample temperature over the nominal value was measured (using a thermocouple) to be +3°C. The temperature quoted in all data presented is the 'true' value not the nominal value. We report here the first application of the new cell: to the polymerization of tubulin into microtubules. This is not the first time that flow dichroism has been used to investigate tubulin polymerization. Near UV flow dichroism of brain microtubules was first reported by Taniguchi and Kuriyama in 1978,[6] and this technique was later revisited by Nordh et al.[5] The wavelength spectra collected by these two groups for tubulin microtubules differ somewhat in shape.

MATERIALS AND METHODS

Design and Construction of Thermostatted Micro Volume Couette LD *Cell*

The design and construction of a micro volume Couette *LD* cell by Crystal Precision Optics, Rugby, U.K., has been described previously.[15,22,23] In the cell used in this work, a quartz capillary of ~5 mm outer diameter, and 3.0 mm inner diameter rotates about a fixed rod of 2.5 mm outer diameter. The capillaries and rods were supplied by Enterprise-Q, Manchester, U.K. Araldite Rapid® was used to seal one end of the quartz capillary. A 2.54 cm diameter × 10 cm focal length uncoated lens (supplier Edmund Optics®) was placed 10 cm in front of the center of the rod to focus the light beam onto the sample. A second uncoated lens (2.54 cm diameter × 5 cm focal length) was placed after the sample to focus all photons onto the photomultiplier tube. A separate unit holds the temperature controller. The quartz components had low intrinsic absorbance and small intrinsic *LD* signals which were subtracted from the data presented below. The operating temperature is between room temperature and ~60°C. The temperature can be maintained to ±0.2% of the set temperature.

Application of Thermostatted Microvolume Couette LD *Cell*

A Jasco J-715 *CD* spectropolarimeter with large sample compartment that has been adapted for *LD* measurements was used for all *LD* experiments. Samples were individually prepared with volumes of 40 μL placed into the capillary using a Gilson air displacement pipette (P100). A 4 V supply was continuously applied to the motor of the Couette cell by an EP-603 (0–30 V) power supply, adapted to allow more precise measurements (to two decimal places) of applied voltage by the addition of a 10-turn potentiometer.

Polymerization of Tubulin

Tubulin protein from bovine brain (Cytoskeleton, Tebu-Bio Ltd., U.K.) was prepared using the supplied buffers and protocol. The general tubulin buffer contains 80 mM Na-Pipes (Piperazine-1,4-bis(2-ethanesulfonic acid)) pH 6.9, 1 mM $MgCl_2$, 1 mM EGTA and 4.8% v/v glycerol. GTP was added to a final concentration of 1 mM. 40 μL of sample was transferred to a capillary and incubated at 37°C within the micro volume Couette *LD* cell. The final concentration of microtubules was ~28 μM. This enabled real-time monitoring of the polymerization of tubulin microtubules using the interval scan measurement program within the Jasco software. Full-wavelength spectra from 350 to 200 nm were collected every minute at 200 nm min^{-1} and data pitch 0.5 nm, with a response of 0.5 sec. A bandwidth of 2 nm was used.

Substrates Binding to Tubulin and Tubulin Microtubules

Paclitaxel (Taxol™) from Taxus yannanensis (Sigma, Poole, U.K.) was prepared in DMSO (99.9% purity, Sigma) and diluted to 200 μM using the general tubulin buffer (see above). The taxol solution was stored at 37°C until required. Colchicine and vincristine were both purchased from Sigma, Poole, U.K. and reconstituted using high-purity water to working concentrations of 280 and 2.8 μM, respectively.

The effects of the ligands on tubulin were probed by allowing tubulin to polymerize for 40 min in the thermostatted micro volume Couette *LD* cell. After 40 min the rotation of the cell was stopped and a single substrate was added to a final tubulin microtubule concentration of ~26 μM (Taxol ~ 18 μM; Colchicine 26 μM and Vincristine 2.6 μM). A constant temperature of 37°C was maintained throughout. Full-wavelength spectra were collected from 350 to 200 nm every minute at 200 nm min^{-1}, data pitch 0.5 nm and with a response of 0.5 sec. A bandwidth of 2 nm was used.

Light-Scattering Correction

The microtubules formed give rise to some light scattering, even with the post sample focusing lens. This is a problem that has previously been observed by Taniguchi and Kuriyama, and Nordh et al.[5,6] Correction for light scattering in the spectra shown here uses the method given by Nordh et al. [eq. (3) and Fig. 3], where LD^T is the background turbidity dichroism, α is constant and k is a constant that has been shown to generally relate to the unpolarised turbidity (k usually equals a value between 2.8 and 3.5).

$$LD^T(\lambda) = \alpha\lambda^{-k} \quad (3)$$

Reduced Linear Dichroism

The reduced linear dichroism was calculated by converting the high-tension voltage recorded by the J-715 to absorbance within the Jasco spectra analysis software using a water *LD* spectrum as a baseline, then calculating the *LD* divided by absorbance [eq. (2)].

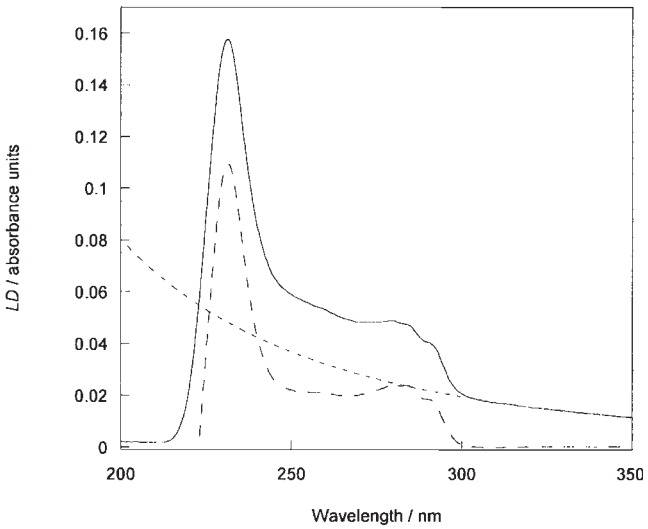


Fig. 3. Method of light-scattering correction when applied to a polymerized tubulin *LD* spectrum.

RESULTS

Polymerization and Depolymerization of Tubulin

A temperature of 37°C is required for the polymerization of tubulin heterodimers into protofilaments, therefore the thermostatted micro volume Couette *LD* cell proved to be ideally suited for the linear dichroism analysis of this system. No *LD* signal is observed when tubulin is in its monomeric/dimeric state. During polymerization, as the microtubule is forming, an increase in net *LD* signal intensity occurs. The *LD* signal is not only dependent on the length of the microtubule but also its flexibility—as the fiber becomes 'stiffer' a further increase in net *LD* is observed. Figure 4(a) shows the full-wavelength spectra of tubulin (~28 μM) at different time points throughout polymerization. Figure 4(b) shows the same wavelength range, but for the depolymerization time points. Both plots have been *LD* baseline corrected, but not corrected for light scattering. It is observed that as the microtubules form there is a large increase in the amount of background light scattering, as shown by a nonflat baseline where there is no absorbance. Figure 4(c) shows kinetics plots at two wavelengths, 250 and 280 nm, of tubulin polymerization and depolymerization. It can be seen that there is a rapid increase in *LD* signal intensity during the first 10 min, followed by a plateau for ~40 min. It is during this plateau that the effects of other molecules on the microtubules may be investigated. There is a slow decrease in *LD* signal as the microtubule depolymerises. After five hours the microtubules have completely depolymerised, as seen by a zero *LD* signal (and no light scattering).

The data in Figures 5 and 6 are not reliable at the low-wavelength end of the spectra. The low-wavelength data have been included to illustrate that, with such highly

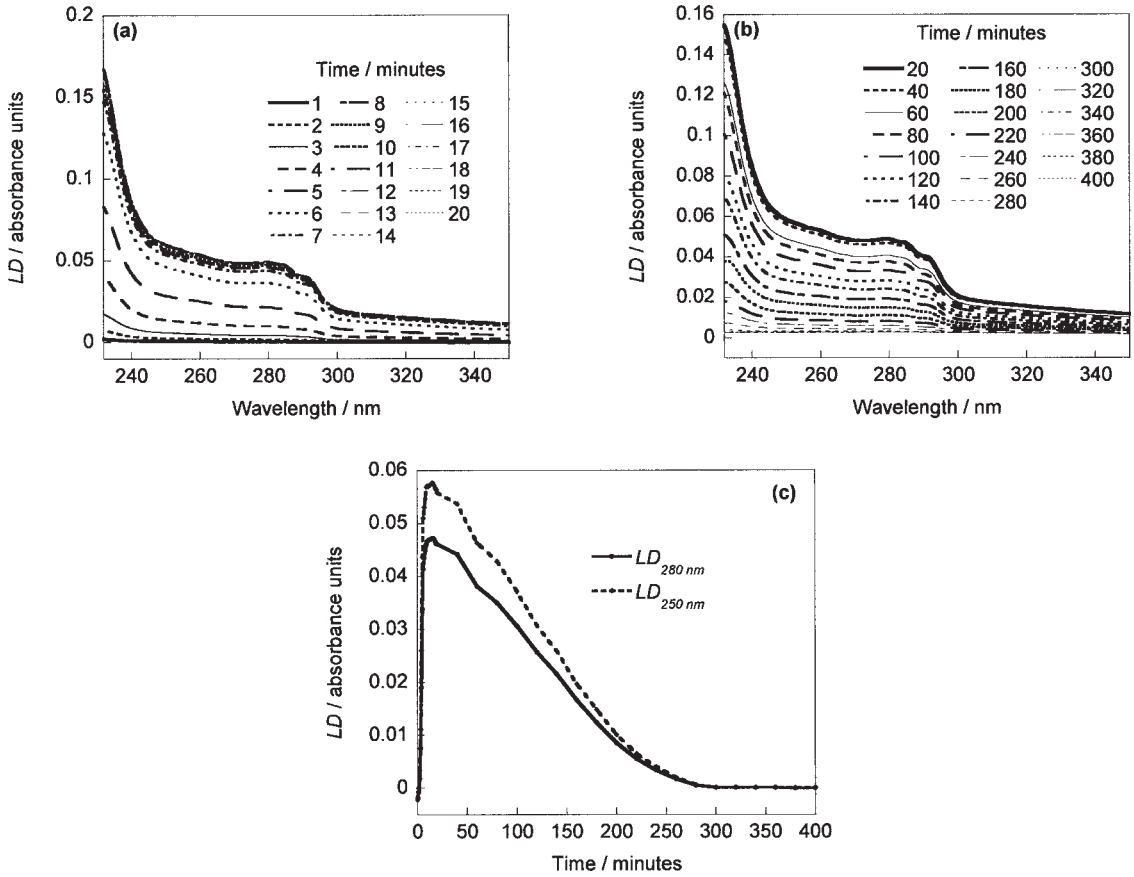


Fig. 4. Baseline corrected *LD* of tubulin microtubules (~28 μM): (a) full-wavelength spectra at different time points throughout polymerization; (b) full-wavelength spectra at different time points throughout depolymerization; (c) *LD* as a function of time at 250 and 280 nm.

light-scattering samples, spectra can look good even when they are not. The reliable high-tension voltage (HTV) range on the spectropolarimeter seems to be less with such highly scattering samples than for nonscattering samples. We have observed the same artifact with other fibrous proteins, for example, FtsZ, actin, and collagen. The spectra of such samples should be checked by diluting the sample. If the wavelength of a signal maximum (positive or negative) moves to lower wavelengths (as in Fig. 5) when the sample is diluted and/or the signal does not follow the Beer–Lambert Law, then the data should be discarded. Despite this problem and our inability to dilute the sample below a 'critical polymerization concentration' we can use our data to probe the protein backbone orientation, since the data above about 237 nm in the near UV of Figures 5 and 6 are true *LD* signals for all concentrations shown. This wavelength region includes the tail of the $n \rightarrow \pi^*$ transition of α-helices at ~220 nm (showing as a positive *LD* signal at 240 nm), so we can actually probe the orientation of the helices. However, other structural motifs such as β-sheets have little *LD* signal in this region, so we cannot investigate them.

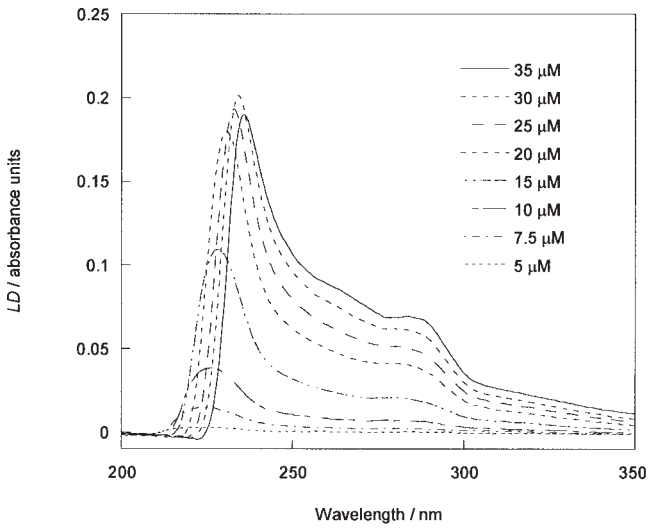


Fig. 5. Investigation of *LD* signal as a function of concentration of tubulin. Dilutions of a stock concentration of 35 μM were made using high-purity water.

LD *Spectra Interpretation*

Figure 6 shows a baseline and light-scattering corrected *LD* spectrum of tubulin after 20 min of polymerization of the heterodimers into microtubules. The net *LD* signals are positive for the accessible far UV. As discussed previously, even though the 232.5 nm *LD* maximum is not real, we do know that the *LD* signal for the

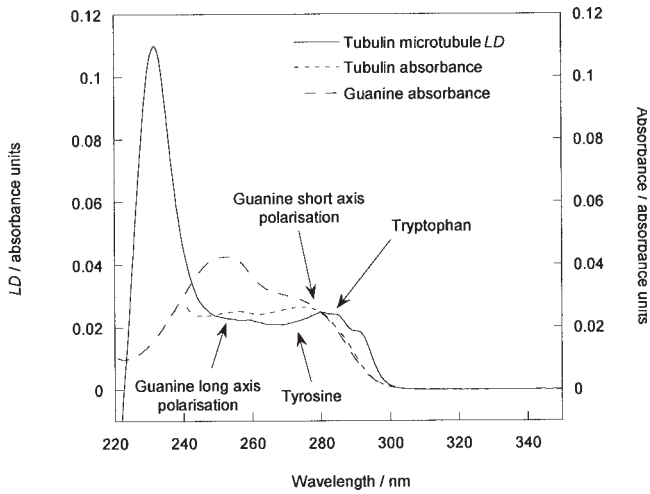


Fig. 6. Baseline and light-scattering corrected *LD* spectrum of tubulin microtubules alone (28 μM, solid line) using the method given by Nordh et al.,[5] the absorbance spectra of tubulin measured in a 1 cm path length cuvette (without GTP, dotted line) and the absorbance spectrum of guanine[13] (dashed line) both rescaled at 280 nm so as to equal the *LD* magnitude of tubulin microtubules.

$n \rightarrow \pi^*$ transition (which is centered at ~220 nm) is positive. (The obvious approach of diluting the sample for these experiments is not feasible as tubulin will not polymerize.) This indicates that the $n \rightarrow \pi^*$ transitions of α-helices in tubulin are aligned more parallel than perpendicular to the orientation axis. Since the $n \rightarrow \pi^*$ transition moment is approximately perpendicular to the α-helix axis,[24] the α-helices may be concluded to be on average aligned more perpendicular than parallel to the orientation axis.[21]

The near UV *LD* also shows a net positive *LD* signal, which tells us that the dominant transitions are oriented more parallel than perpendicular to the orientation axis. By overlaying the absorbance spectra of tubulin monomers and of guanosine on the *LD* spectrum of polymeric tubulin (scaled to be equal to the *LD* at 280 nm since we cannot determine the orientation factor *S*, see Fig. 6), it is possible to identify which chromophore(s) dominate(s) the spectrum at a given wavelength. Guanine in GTP/GDP has two transitions in the near UV: a long axis-polarized transition at 248 nm, and a short axis-polarized transition at 279 nm.[25,26] Tyrosine (4% of amino acids) and tryptophan (<1% of amino acids) have transitions with maxima at 274 and 280 nm, respectively, and phenylalanine (5% of amino acids) has its maximum at ~260 nm. The 290 nm region is thus dominated by the tail of the tryptophan band. Tryptophan dominance has been previously observed with flow orientation of the tobacco mosaic virus, which only has one tryptophan.[27,28] The LD^r (reduced linear dichroism) makes the contributions from the different transitions more evident (Fig. 7). LD^r removes concentration, absorbance intensity, and path-length factors. The long axis-polarized transition of guanine dominates the 250–260 nm region (phenylalanine $\varepsilon_{258\ nm}$ is very small); the fact that the LD^r is a flat line indicates there is a single transition polarization. The increase in LD^r at wavelengths above 260 nm shows that the transitions occurring here are oriented more parallel to the orientation axis than is the long axis of guanine. No flat LD^r region is observed above 260 nm, suggesting that there are a number of overlapping transitions of different polarizations. These include a short axis-polarized transition of guanine, and transitions of tyrosine and of tryptophan. As observed from Figure 7, the 290 nm region, which is dominated by tryptophan residues, is oriented closest to the fiber axis (as seen by the maximum positive LD^r). The short axis-polarized transitions of guanine and the tyrosine transition are oriented intermediate between the guanine long axis-polarized transition and the average tryptophan long axis-polarized transition.

Full-wavelength *LD* spectra were collected at regular intervals throughout tubulin polymerization and depolymerization. Light-scattering corrected spectra ($k = 3.5$) are shown in Figure 8(a). Figure 8(b) shows the LD^r at the same time points as in Figure 8(a). The absorbance data also showed significant light scattering; the scattering component had been removed from the plots of Figure 8 using the same method as that for the *LD* (using $k = 3.5$). It is interesting to note that the shape of the spectra changes between 240 and 280 nm as time progresses; however, the shape of the 285–300 nm region (transition of the tryptophan residues only) remains unchanged. It is not possible to determine from the LD^r whether the short axis of guanine is reorienting during polymerization and depolymerization as tyrosine also absorbs in this region. During the first 10 min, the LD^r increases in intensity showing that all chromophores are becoming oriented as the fibers form. The tryptophan region increases in LD^r more than the others. This shows that the tryptophan long axis is becoming oriented more parallel to the fiber axis than the other residues (guanine and tyrosine) during polymerization. LD^r values at the

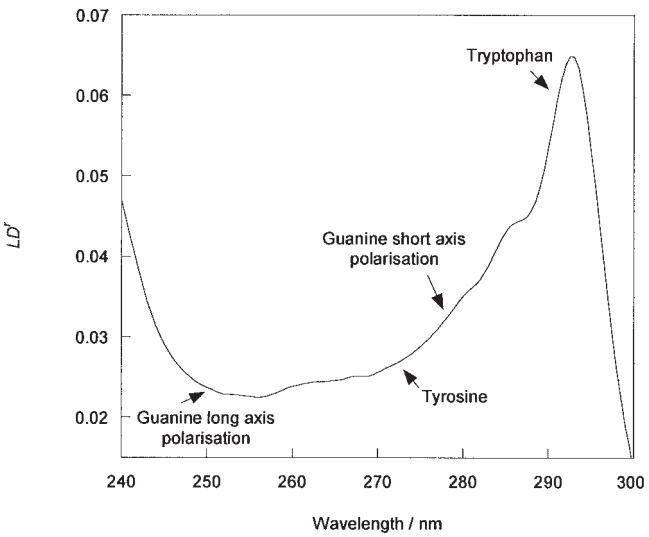


Fig. 7. LD^r (linear dichroism/absorbance) of tubulin microtubules alone (~28 μM). Absorbance calculated by converting the HT voltage to absorbance within the Jasco spectra analysis software (using a water *LD* spectrum as a baseline).

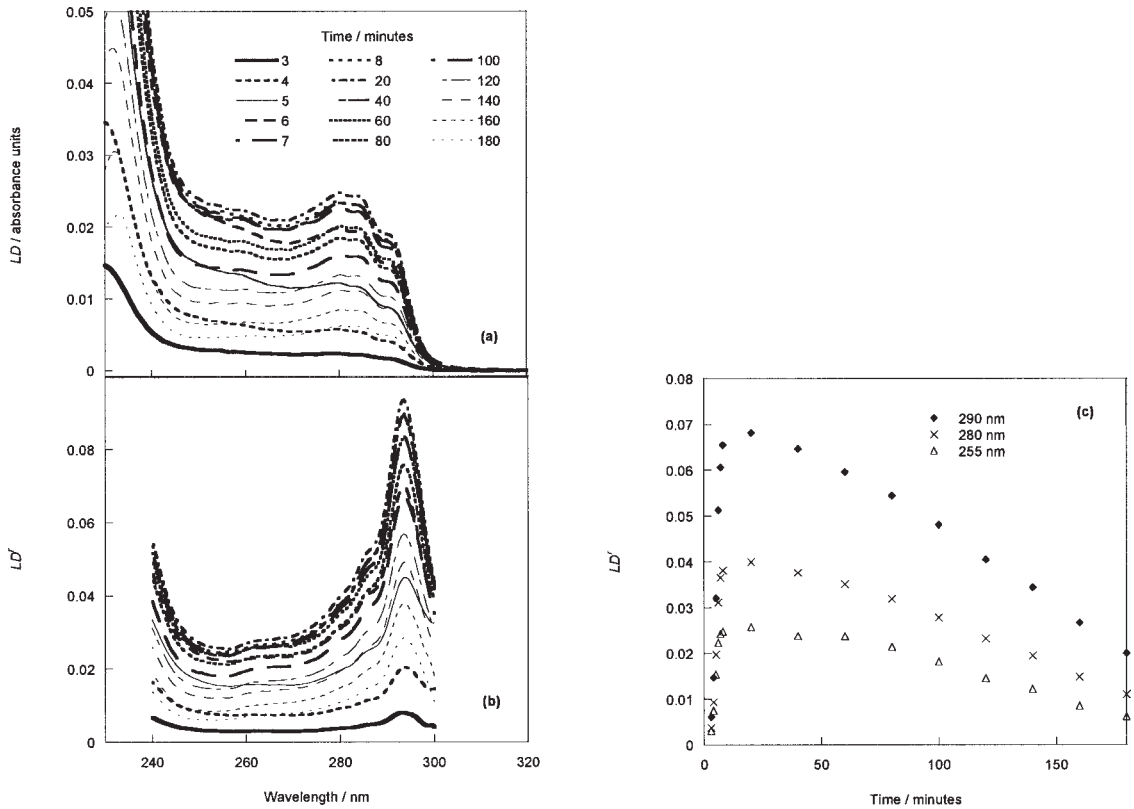


Fig. 8. (a) Baseline and light-scattering corrected *LD* spectrum of tubulin microtubules (28 μM) at different time points throughout polymerization and depolymerization; (b) LD^r (reduced linear dichroism) at the same time points as shown in (a); (c) LD^r at 255, 280, and 290 nm.

wavelengths 290 nm (trytophan); 280 nm (tyrosine, tryptophan and short axis polarization of guanine) and 255 nm (long axis polarization of guanine) are shown as a function of time in Figure 8(c). The kinetic profiles indicate an initially rapid (first 10 min) process, followed by a short plateau, before a steady decrease is observed. The first phase (1–5 min) of the rapid magnitude increase corresponds to the initial polymerization of heterodimers into protofilaments. The subsequent reorientation is consistent with the increasing curvature of protofilaments that occurs during microtubule assembly,[29] which our data show occurs with the increase in the parallel orientation of tryptophan residues, and an increase in stiffness (overall orientation) upon forming a microtubule.

Depolymerization occurs all the time once a polymeric unit has been formed. Initially, a 'treadmilling' effect is observed when new heterodimers are added to one end of the polymer, as others dissociate from the other end. This happens until all the GTP has been hydrolyzed, at which stage it is not possible to add any new heterodimers; therefore, the polymeric unit undergoes net depolymerization and shortens in length. Consistent with this, a decrease in the LD^r signal intensity is observed after 50 min—as the polymer length decreases and increasing numbers of unbound monomers/dimers contribute to the absorbance but not the *LD*. The shape of the *LD* spectrum does not change during this phase, indicating that the chromophoric units of the polymer remain in the same orientations during this process.

Substrates Binding to Tubulin Microtubules

Tubulin microtubules were prepared as previously described. After 40 min, the rotation of the cell was stopped, a substrate was added to the capillary, and rotation was restarted.

Taxol™

Taxol™ (Fig. 9) is known to enhance and stabilize microtubule assembly. The *LD* showed this with an initial small increase in the *LD* signal intensity due to the concentration of formed microtubules increasing and a decrease in the rate of *LD* signal decay. Even after 15 h an *LD* signal was still measurable in contrast to the situation with Taxol-free polymers. The LD^r at 900 min reveals Taxol-induced changes in the orientation of all aromatic chromophores (Fig. 10). It has been reported that Taxol™ bound to polymerized tubulin does not modify the conformation of the bound nucleotide;[30] unfortunately, from the *LD* data analysis perspective, a Taxol™ transition overlaps with the guanine contributions in this region, making it difficult to interpret the data without a theoretical model of the electronic spectroscopy of this system. The 250 nm transition, which is partly due to the long axis polarization of guanine and a transition of Taxol™, has a negative *LD* signal, indicating that the polarizations responsible for this are aligned on average more perpendicular than parallel to the orientation axis. Work is in progress for the quantitative analysis of *LD* in

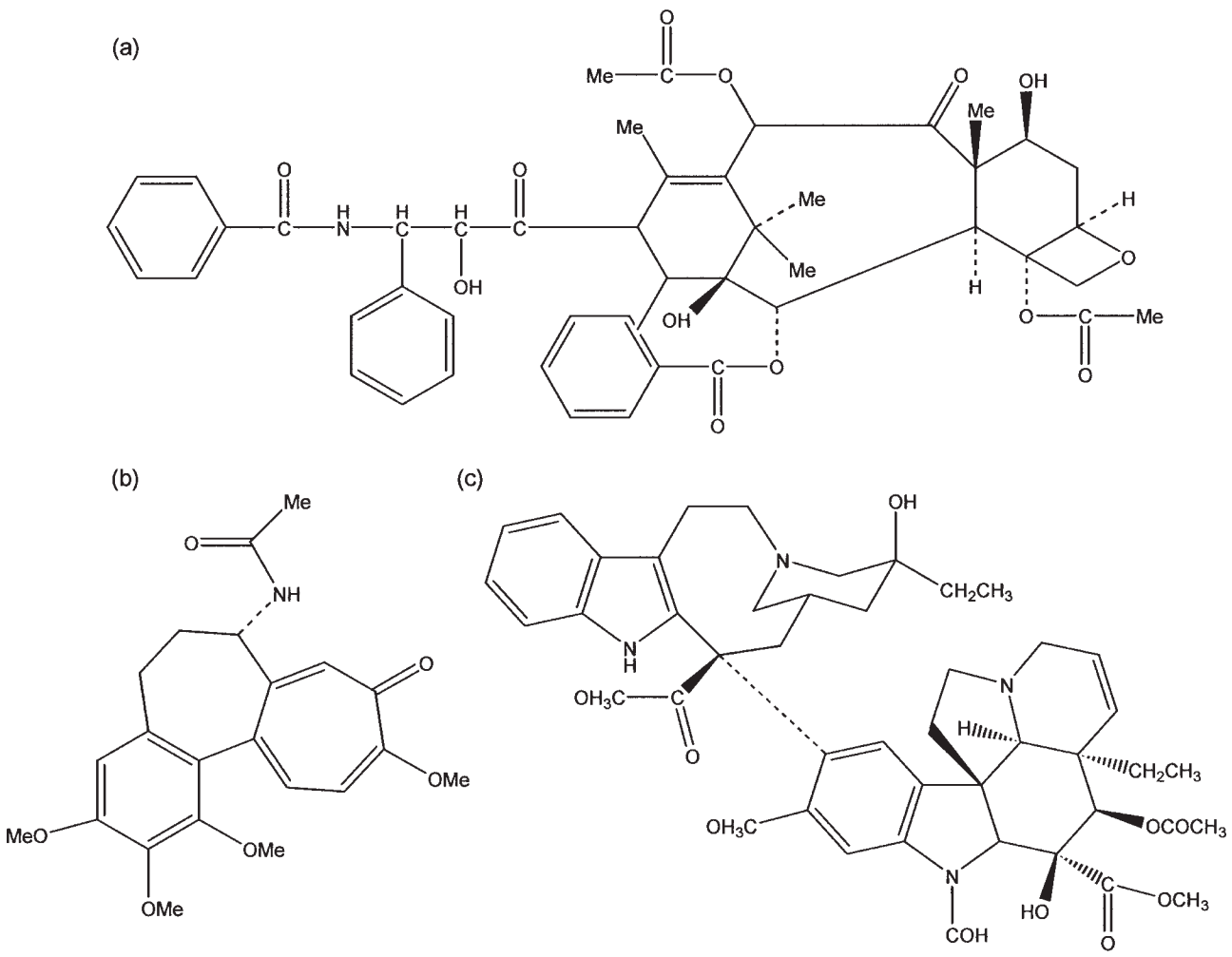


Fig. 9. (a) Taxol; (b) colchicine; (c) vincristine.

terms of spectroscopy of the component parts (e.g., Taxol™, tyrosine, guanine, and so forth).

Colchicine

Colchicine (Fig. 9) binds to tubulin heterodimers. When a colchicine-bound heterodimer is incorporated into the tubulin microtubule, it prevents any further addition of tubulin heterodimers. Consequently, a rapid net depolymerization of the microtubules occurs. This is shown in *LD* by a rapid decrease in the *LD* signal intensity of tubulin after colchicine was added to the polymerized microtubules (Fig. 11).

Once a light-scattering correction is taken into account, small differences in the near UV *LD* between tubulin microtubules alone and tubulin microtubules–colchicine are apparent (Fig. 12). The *LD* spectrum of tubulin microtubules–colchicine has been rescaled at 280 nm in order to allow easier comparison against tubulin microtubules alone. The rescaled *LD* spectrum of tubulin microtubules–colchicine shows a shape change in the *LD* from 245 to 280 nm. When rescaled at 280 nm, this region is smaller in magnitude for tubulin microtubules–colchicine than tubulin microtubules alone. The corresponding LD^r spectra do not show anything conclusive other than a decrease in intensity, which is in part due to the decrease in protein concentration in the polymer versus in the sample as a whole not being taken into account in the overall concentration. Colchicine has an absorbance at 250 nm; therefore, one must ask whether the effects observed in the *LD* are due to colchicine being oriented on the microtubules or due to changes within the microtubule itself. Only one colchicine molecule would be bound per microtubule (as the addition of a tubulin–colchicine subunit hinders the addition of further subunits); therefore, in comparative terms, the *LD* is unlikely to be able to probe the binding of colchicine because it is not oriented effectively (being located at the end of a microtubule) and/or it is not at a sufficient orientation. Consistent with the expected low concentration of bound colchicine, there is no *LD* signal at 350 nm (where the absorbance of the 26 μM colchicines solution is ~0.4, data not shown) and no peak at 250 nm (where there is a fairly sharp absorbance of a 26 μM solution of colchicines of magnitude ~0.8, data not shown), so either both the 350 and 250 nm transitions are fortuitously oriented at or close to the magic angle on the tubulin, or the bound col-

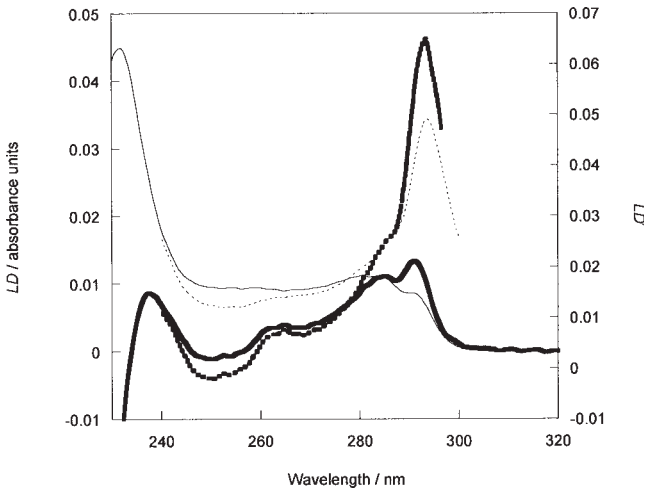


Fig. 10. *LD* spectra (solid lines) and LD^r (dashed lines) of tubulin microtubules alone (28 μM) at time point 140 min (thin solid and dashed lines) and tubulin microtubules (26 μM)-Taxol™ (18 μM) at 900 min (thick solid and dashed lines) with a correction for light scattering for both *LD* and absorbance.

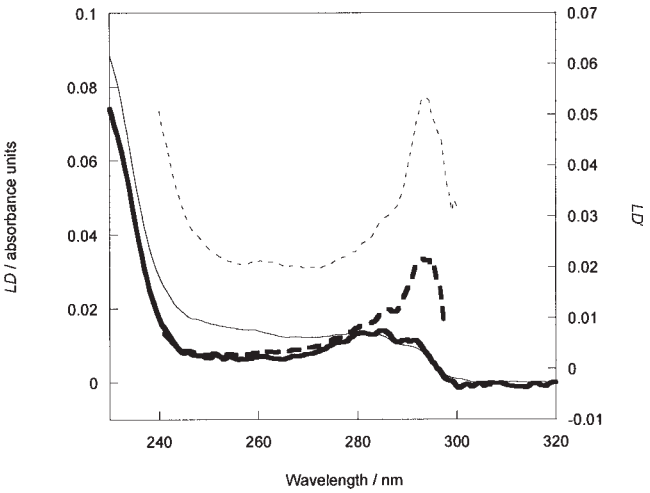


Fig. 12. *LD* spectra (solid lines) and LD^r (dashed lines) of tubulin microtubules alone (28 μM) at time point 140 min (thin lines) and tubulin microtubules–colchicine (both 26 μM) at 41 min (thick lines) (rescaled at 280 nm) with a correction for light scattering for both *LD* and absorbance.

chicine is too dilute to be detected and what we are measuring in the UV region is indeed the effect on the tubulin microtubule structure due to colchicine.

Vincristine

It has previously been reported that vincristine (Fig. 9) promotes the curling and aggregation of tubulin microtubules.[10–12] Figure 13 shows the comparative *LD* spectra of tubulin microtubules alone and the first measurement (at the one minute) after addition of vincristine. It can be seen that there is an immediate decrease in the *LD* signal upon addition of vincristine; however, the end product *LD* spectrum is different from that when the substrate was colchicine. For vincristine, a nonflat baseline is observed (see inset of Fig. 13) for the final product, indicating it is not simply disassembled to monomeric or dimeric units in accord with literature data. It was not possible with our cell to monitor the kinetics of depolymerization—a form of stopped flow apparatus would be required to monitor this due to the speed of disassembly of the microtubule.

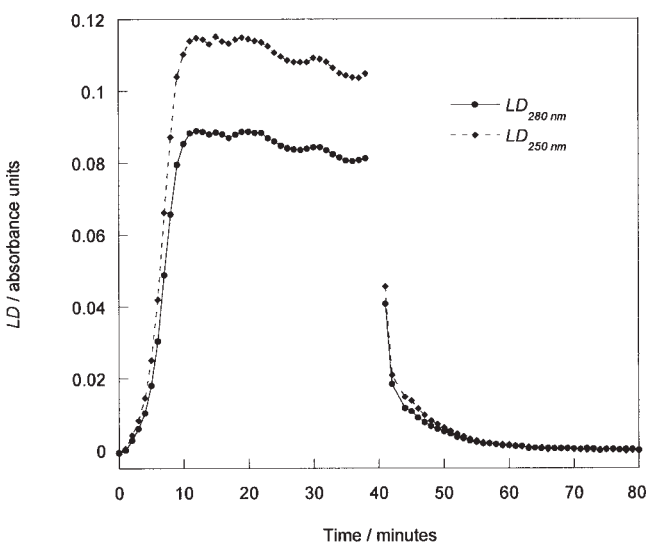


Fig. 11. *LD* kinetics analysis of tubulin microtubules (~26 μM) and colchicine (26 μM).

DISCUSSION AND CONCLUSION

A micro volume Couette *LD* cell has been designed and developed that has the capability to be thermostatically controlled. The small sample volume (20–40 μL) required by the cell has widened the application of *LD* to

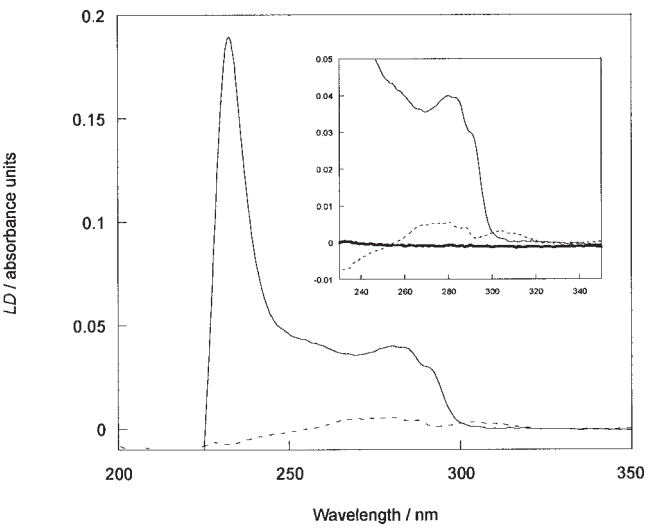


Fig. 13. Comparison of *LD* spectra of tubulin microtubules alone (28 μM) (solid lines) and tubulin microtubules–vincristine (both 2.6 μM) (dashed lines) immediately after addition of vincristine at 41 min with light scattering correction. Inset also includes the *LD* spectrum of the end product of tubulin microtubules–vincristine (both 26 μM) at 70 min (thick solid line).

be able to include extensive studies of fibrous proteins. In this work, being able to hold the polymerising solutions at 37° for 15 h was important for the taxol-stabilized polymers. This new element of temperature regulation enables the analysis of temperature sensitive polymeric material. Though a thermostatted Couette *LD* cell has been used to investigate the *LD* of tubulin previously,[5] the micro volume Couette *LD* cell significantly reduces the volume required and enables a longer wavelength range to be reported so that the it is possible now to probe the tail of the backbone region. Previous analysis by Taniguchi and Kuriyama, and Nordh et al. concluded that the electronic transition moments of both nucleotides and protein aromatic groups are oriented parallel to the orientation axis.[5,6] The *LD* spectra reported here are qualitatively in accord with those previously recorded in terms of spectral shape. However, because we were able to collect wavelength scans throughout the polymerization and depolymerization cycles, and also calculate the LD^r, we have shown that there are orientation changes of the guanine and tryptophans during polymerization. During polymerization, the trytophan residues increase in orientation more than other residues and so become on average the most parallel to the orientation axis. Due to the dominant *LD* signal from tryptophan overlapping the long axis-polarized transition of guanine, it is difficult to determine the orientation of guanine; however, it shows little change during polymerization. Work is in progress to use the data reported here as input to calculations of *LD* spectra to enable such information to be extracted. We have found the nucleotides of other fibrous proteins such as FtsZ[31] to rearrange during the different processes the fiber undergoes.

Differences between the *LD* spectra reported by Taniguchi and Kuriyama, and those of Nordh et al. (the relative intensity of the 260 nm *LD* signal compared to that at 280 nm is significantly less for Taniguchi and Kuriyamas' data) can be understood using the results reported here as we see precisely this change for microtubule spectra collected at different time points after initiation of polymerization. The reorientation of the tryptophan is responsible for the difference.

The effects of the ligands Taxol™, colchicine, and vincristine on tubulin have been probed by *LD* in this work, giving an indication of their structural effects. Taxol™, as expected, stabilizes the tubulin microtubules once they are formed and also cause a large decrease in the *LD* signal intensity between 240 and 280 nm, which is where an absorbance for Taxol™ lies. This implies that either Taxol™ is oriented perpendicular to the orientation axis or Taxol™ is affecting the orientation of the long axis of guanine. At present it is not possible to differentiate these two options, though work is in progress to theoretically calculate the component parts of the *LD* spectrum. Colchicine has a large effect on the tubulin polymers, in that an immediate loss of orientation of the tubulin microtubule is observed. The *LD* spectra of tubulin microtubules in the presence of colchicine show changes in the shape of the *LD* at 250 nm, where the ligand absorbs. Although it is difficult to determine conclusively whether the *LD* changes are probing the binding of colchicine and/or its effect on the microtubule structure, as there is no sharp *LD* signal at 250 nm and 350 nm, we conclude the *LD* is showing changes in the fiber structure as a function of time. Vincristine immediately (on the timescale of our experiment where it takes 10 sec to add the ligand) disrupted the microtubule and caused loss of structure. The kinetics are too fast to be monitored using this method; therefore, some form of stopped flow *LD* apparatus would be required, which is not available at present.

In conclusion, the thermostatted micro volume Couette *LD* cell offers the potential for temperature-sensitive analysis of fibrous proteins, as demonstrated by the analysis of tubulin microtubules. It has been possible to probe further the structural reorientations of tubulin chromophores throughout polymerization and the binding of ligands. This and further analysis using the technique of *LD* could yield new information for molecular cell biologists.

ACKNOWLEDGMENTS

The collaboration of Crystal Precision Optics, Rugby UK has been a key factor in the design and building of the thermostatted microvolume Couette *LD* cell and is gratefully acknowledged. R. Marrington also thanks Syngenta for a Ph.D. studentship. Funding from the Engineering and Physical Sciences Research Council is gratefully acknowledged.

LITERATURE CITED

1. Luduena RF, Shooter EM, and Wilson L. Structure of the tubulin dimer. J Bio Chem 1977;252:7006–7014.
2. Desai A and Mitchison TJ. Microtubule polymerisation dynamics. Ann Rev Cell Developmental Bio 1997;13:83–117.
3. Maccioni R and Seeds NW. Stoichiometry of GTP hydrolysis and tubulin polymerisation. Proc Natl Academy Sciences USA 1977;74: 462–466.
4. Walter A, Gast K, Vater W, and Zirwer D. Measuring of the *in vitro* assembly of microtubules by dynamic light scattering. Acta Histochemica 1992;41:29–35.
5. Nordh J, Deinum J, and Norden B. Flow orientation of brain microtubules studied by linear dichroism. Europ Biophys J 1986;14:113–122.
6. Taniguchi M and Kuriyama R. Ultraviolet flow dichroism of brain microtubule. Biochimica Et Biophysica Acta 1978;533:538–541.
7. Jordan MA and Wilson L. Use of drugs to study role of microtubule assembly dynamics in living cells. Methods in Enzymology 1998;298: 252–276.
8. Parness J and Horwitz SB. Taxol binds to polymerized tubulin *in vitro*. J Cell Bio 1981;91:479–487.
9. Skoufias DA and Wilson L. Mechanism of inhibition of microtubule polymerisation by colchicine: inhibitory potencies of unliganded Colchicine and Tubulin–Colchicine Complexes. Biochem 1992;31: 738–746.
10. Himes RH. Interactions of the catharanthus (vinca) alkaloids with tubulin and microtubules. Pharmacology therapy? 1991;51:257–267.
11. Lobert S, Vulevic B, and Correia JJ. Interaction of vinca alkaloids with tubulin: a comparison of vinblastine, vincristine, and vinorelbine. Biochem 1996;35:6806–6814.
12. Prakash V and Timasheff SN. The interaction of vincristine with calf brain tubulin. J Biol Chem 1983;258:1689–1697.
13. Rodger A and Nordén B. Circular dichroism and linear dichroism. Oxford: Oxford University Press; 1997. p 150.

14. Taylor GI. Fluid friction between rotating cylinders II. Distribution of velocity between concentric cylinders when outer one is rotating and inner one is at rest. Proc Royal Society of London Series A 1936; 157:565–578.
15. Marrington R, Dafforn TR, Halsall DJ, MacDonald JI, Hicks M, and Rodger A. Validation of new microvolume Couette flow linear dichroism cells. The Analyst (in press).
16. Wada A and Kozawa S. Instrument for the studies of differential flow dichroism of polymer solutions. J Polymer Sci Part A 1964;2:853–864.
17. Nordén B. Applications of linear dchroism spectroscopy. applied Spectroscopy Rev 1978;14:157–248.
18. Higashi S, Kasai M, Oosawa F, and Wada A. Ultraviolet dichroism of F-actin oriented by flow. J Molecular Bio 1963;7:421–430.
19. Oosawa F, Maeda Y, Fujime S, Ishiwata S, Yanagida T, and Taniguchi M. Dynamic characteristics of F-Actin and thin filaments *in vivo* and *in vitro*. J Mechanochem Cell Motility 1977;4:63–78.
20. Hofricheter J and Eaton W. Linear Dichroism of biological chromophores. Ann Rev Biophys Bioengg 1976;5:511–560.
21. Dafforn TR, Rajendra J, Halsall DJ, Serpell LC, and Rodger A. Protein fibre linear dichroism for structure determination and kinetics in a low-volume, low-wavelength Couette flow cell. Biophys J 2004;86: 404–410.
22. Marrington R, Dafforn TR, Halsall DJ, and Rodger A. Micro-volume Couette flow sample orientation for absorbance and fluorescence linear dichroism. Biophys J 2004;87:2002–2012.
23. Marrington R, Dafforn TR, Halsall DJ, Hicks M, and Rodger A. Validation of new microvolume Couette flow linear dichroism cells. Analyst 2005;130:1608–1616.
24. Rodger A, Rajendra J, Marrington R, Ardhammar M, Nordén B, Hirst JD, Gilbert ATB, Dafforn TR, Halsall DJ, Woolhead CA, and others. Flow oriented linear dichroism to probe protein orientation in membrane environments. Phys Chem Chem Phys 2002;4:4051–4057.
25. Clark LB. Electronic Spectra of crystalline 9-ethylguanine and guanine hydrochloride. J Amer Chemical Soc 1977;99:3934–3938.
26. Fülscher MP, Serrano-Andrés L, and Roos BO. A theoretical study of the electronic spectra of adenine and guanine. J Amer Chemical Soc 1997;119:6168–6176.
27. Taniguchi M, Yamaguchi A, and Taniguchi T. Flow dichroic spectra of tobacco mosaic virus and their protein assemblies. Biochimica Et Biophysica Acta 1971;251:164–171.
28. Rees MW and Short MN. Variations in the composition of two strains of tobacco mosaic virus in relation to their host. Virology 1965;26:596–602.
29. Löwe J, van den Ent F, and Amos LA. Molecules of the bacterial cytoskeleton. Ann Rev Biomolecular Structure 2004;33:177–198.
30. Carlier MF and Pantaloni D. Taxol effect on tubulin polymerization and associated guanosine 5′-triphosphate hydrolysis. Biochem 1983; 22:4814–4822.
31. Marrington R, Small E, Rodger A, Dafforn TR, and Addinall S. FtsZ fibre bundling is triggered by a calcium-induced conformational change in bound GTP. J Biol Chem 2004;47:48821–48829.

Restriction Enzyme Kinetics Monitored by UV Linear Dichroism†

Matthew R. Hicks,*,‡,§ Alison Rodger,‡ Christopher M. Thomas,§ Sarah M. Batt,§ and Timothy R. Dafforn§

Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K., and Department of Biological Sciences, University of Birmingham, Birmingham B15 2TT, U.K.

Received January 26, 2006; Revised Manuscript Received April 24, 2006

ABSTRACT: The use of linear dichroism (LD) spectroscopy for biological applications has been brought to the forefront recently by our development of thermostated microvolume Couette cells. We present a method for following the digestion of DNA by restriction endonucleases in real time without the use of any extrinsic dyes or labels. This is accomplished using linear dichroism spectroscopy (the differential absorbance of light polarized parallel and perpendicular to the sample orientation axis). The differential absorbance signal depends on the degree of alignment of the molecules. In this case the DNA is aligned by Couette flow (flowing the solution in the annular gap between two concentric cylinders), and we monitor the increase in alignment upon linearization of a circular DNA molecule. In addition, we observe a decrease in alignment upon further digestion and subsequent shortening of the DNA. Ten enzymes were investigated: seven enzymes with a single cut site (*Eco*RI, *Kpn*I, *Nde*I, *Not*I, *Nru*I, *Sma*I, *Xba*I), two enzymes with two cut sites (*Bst*Z17I, *Eag*I), and one enzyme with no cut site (*Cla*I). LD, as implemented in this new assay, is broadly applicable across a wide range of DNA-modifying enzymes and compounds and, as such, is a useful addition to the toolbox of biological characterization.

The type II restriction endonucleases are a family of enzymes that recognize specific DNA sequences and cleave at fixed positions at or close to that sequence to produce 5′-phosphates and 3′-hydroxyls. They may act as monomers, dimers, or tetramers (*1*). Restriction endonucleases evolved to protect prokaryotic organisms from invasion by foreign DNA; their own DNA is protected by specific DNA methylation patterns (N4 or C5 methylation of cytosine or N6 methylation of adenine) whereas the foreign DNA is not. The foreign DNA is therefore susceptible to cleavage (restriction) by the host enzyme (*2*). The thousands of restriction enzymes that have been identified to date are also extremely important tools for the molecular biologist. They are used in many techniques that require DNA manipulation and have applications in, for example, diagnostics and recombinant protein production (*3*).

To detect and characterize the cleavage of DNA (and thus also monitor the activity of the enzyme), one can use a variety of techniques, the most common of which is to separate the DNA fragments, postincubation, according to molecular size using electrophoresis, usually agarose gel electrophoresis. For time-based measurements each time point requires one lane on a gel, limiting the number of points that can be practically measured. It is more desirable to be able to follow the reaction in real time. There have been a number of approaches developed to follow the cleavage of DNA in real time. Many of these are fluorescence-based and include fluorescence resonance energy transfer (FRET)[1] (*4*), dual-color fluorescence cross-correlation spectroscopy (*5*, *6*), fluorescence anisotropy combined with FRET (*7*), or the use of a fluorophore and a quencher at the ends of a nucleic acid hairpin (*8*). Alternative approaches include the use of immobilized nucleic acids on evanescent wave sensors (*9*, *10*) and ferrocene-modified oligonucleotides coupled with electrochemical detection (*11*). All of these approaches require modification of the substrate with either a label or a means of immobilization on a surface. One can measure restriction enzyme activity using an assay based on the hyperchromic effect; however, the substrate DNA (synthetic oligonucleotides) must be designed such that it is double stranded before cleavage and becomes single stranded after cleavage (*12*).

As an alternative to the methods that require DNA modification or the design of specific oligonucleotides, we proposed that linear dichroism (LD) could be used to monitor the cleavage of unmodified DNA in real time. LD is the difference in absorbance of light polarized parallel and perpendicular to the sample orientation axis. Chromophores in the sample (the DNA bases in this case) absorb the two polarizations of light to different extents depending on their

† Funding from the EPSRC (GR/T09224/01) and BBSRC (REI20503) is gratefully acknowledged. T.R.D. is an MRC Career Development Fellow.

* Corresponding author. Telephone: +44 (0)2476 523293. Fax: +44 (0)2476 524112. E-mail: Matthew.Hicks@warwick.ac.uk.

‡ University of Warwick.

§ University of Birmingham.

[1] Abbreviations: A, absorbance; $A_{\parallel}$, absorbance of light polarized parallel to the sample orientation; $A_{\perp}$, absorbance of light polarized perpendicular to the sample orientation; α_i, angle the transition moment makes with the axis of alignment; bp, base pairs; DTT, dithiothreitol; FRET, fluorescence resonance energy transfer; HT, high tension (voltage); i.d., inside diameter; K_M, Michaelis constant; LB, Luria−Bertani; LD, linear dichroism; LD_{260}, linear dichroism signal at 260 nm; LD^r, reduced linear dichroism signal; o.d., outside diameter; PMT, photomultiplier tube; S, the orientation parameter of the sample; [S], substrate concentration; $t_{1/2}$, half-time for restriction of substrate DNA; TAE, Tris−acetate−EDTA buffer; V, velocity of enzyme reaction; V_{max}, maximum velocity of enzyme reaction.

10.1021/bi0601712 CCC: $33.50
Published on Web 06/27/2006

orientation in the sample. The magnitude of the linear dichroic signal for a chosen transition is given by the equation

$$LD^r = \frac{LD}{A} = \frac{A_{||} - A_{\perp}}{A} = \frac{3S}{2}(3\cos^2 \alpha_i - 1) \qquad (1)$$

where LD^r is the reduced LD (LD divided by the total absorbance of the sample, A), $A_{||}$ is the absorbance of light polarized parallel to the sample orientation, and $A_{\perp}$ is the absorbance of light polarized perpendicular to the sample orientation. The angle the ith transition moment makes with the axis of alignment is denoted by α_i, and S is the orientation parameter of the sample: it equals 1 for perfect alignment and 0 for an isotropic sample. See Rodger and Nordén (*13*) for a more comprehensive review of dichroic spectroscopy.

The extent to which a molecule is oriented by shear flow (its S value) depends on several factors including flow rate, viscosity of the solvent, rigidity/flexibility of the molecule, and the ratio of the long and short axes of the molecule. The LD signal is thus expected to change if the extent of orientation (S) and/or the average base orientation (α_i) changes. We hypothesized that the magnitude of the LD signal would change when supercoiled DNA was linearized and also when the initial product was further cleaved. The change in the LD signal, caused by the changes in S and α_i upon cleavage of the DNA, could thus be used as an assay to follow the enzyme reaction in real time without the need for any dyes or labels. The precise structure of the plasmid is unclear; however, because of the relatively compact nature of circular supercoiled molecules and the nonuniform direction of the bases, it would be reasonable to expect the supercoiled plasmid to give a smaller LD signal than the linearized DNA molecule. Further cleavage of the DNA (if there are multiple restriction sites) will result in shorter linear DNA molecules whose orientation in the flow is expected to be less that that of the longer DNA, resulting in a reduction in measured LD. To investigate this, we monitored the LD signal throughout reactions of DNA with different restriction enzymes.

Here we present some examples of the way in which this potentially generally applicable technique could be used. Seven restriction enzymes were tested for cleavage at their chosen single site of the substrate DNA; an enzyme that does not cut was also tested. Two additional enzymes that cut at two sites each were tested to determine if the two kinetic events could be followed.

The data were analyzed by calculating Michaelis−Menten kinetic parameters from the LD data; an example of this is given for *Eco*RI using the initial rate method. The kinetics of the enzymes that cut twice are more complicated than the single cut ones. We are working on characterizing these reactions by analyzing the complete reaction progress curves numerically in terms of different kinetic models. However, this sort of in-depth analysis is beyond the scope of the present study and will be the subject of subsequent work.

EXPERIMENTAL PROCEDURES

DNA Purification. The plasmid used in this work was a pCDNA3.1 derivative (Invitrogen, Paisley, U.K.) and was 6882 base pairs (bp) in length. XL10 Gold *Escherichia coli* cells (Stratagene, La Jolla, CA) were transformed with plasmid DNA and grown overnight at 37 °C on Luria−Bertani (LB) agar supplemented with 100 μg/mL (270 μM) sodium salt of ampicillin (Sigma-Aldrich Co. Ltd,, Gillingham, U.K.). A single colony was transferred to LB broth and grown overnight at 37 °C. The DNA was purified from the *E. coli* cell pellets using a Concert Maxi-Prep system (Gibco BRL, Gaithersburg, MD) according to the manufacturer's instructions and dissolved in ultrapure water (resistivity 18.2 MΩ·cm). The DNA was stored frozen at −20 °C in small aliquots until required.

Gel Electrophoresis and Restriction Enzyme Reactions. Electrophoresis was performed in 0.7% (w/v) agarose gels in 1× Tris−acetate−EDTA (TAE) buffer (Fisher Scientific, Loughborough, U.K.) (*3*). All samples were run in 1× DNA loading buffer (Bioline, London, U.K.) (*3*). Samples containing 40 ng of DNA from the end point of the restriction enzyme kinetic experiments were run on agarose gels to check for digestion. The enzymes were supplied by the following companies: *Cla*I and *Nru*I (Invitrogen Ltd., Paisley, U.K.); *Bst*Z17I, *Eag*I, *Eco*RI, *Kpn*I, *Nde*I, *Sma*I, and *Xba*I (New England Biolabs U.K., Hitchin, U.K.); *Not*I (Promega, Southampton, U.K.). The cut sites of the enzymes are as follows with the recognition sequence followed by the location of that sequence in our plasmid; the caret indicates the position in the recognition sequence at which the enzyme cuts: *Bst*Z17I (GTA∧TAC: 1831, 4694), *Cla*I (AT∧CGAT: no site), *Eag*I (C∧GGCCG: 933, 3630), *Eco*RI (G∧AATTC: 2415), *Kpn*I (GGTAC∧C: 2450), *Nde*I (CA∧TATG: 489), *Not*I (GC∧GGCCGC: 934), *Nru*I (TCG∧CGA: 212), *Sma*I (CCC∧GGG: 3535), and *Xba*I (T∧CTAGA: 921).

For production of linear DNA (used to assess LD signal as a function of rotation speed and LD signal as a function of concentration), plasmid DNA (10 μg) was digested at 37 °C for 2 h with 50 units of *Eco*RI (2.5 μL) in the buffer supplied with the enzyme (50 mM NaCl, 100 mM Tris-HCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5) in a total volume of 50 μL. Size markers were kilobase DNA markers (Amersham Pharmacia Biotech, Chalfont St. Giles, U.K.). Buffers for other enzymes were as supplied and as follows: *Cla*I (50 mM Tris-HCl, 10 mM $MgCl_2$, pH 8.0), *Nru*I (50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 50 mM KCl, pH 8.0), *Bst*Z17I (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), *Eag*I (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), *Eco*RI (50 mM NaCl, 100 mM Tris-HCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5), *Kpn*I (10 mM Bis-Tris propane HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.0), *Nde*I (50 mM potassium acetate, 20 mM Tris−acetate, 10 mM magnesium acetate, 1 mM DTT, pH 7.9), *Sma*I (20 mM Tris−acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, pH 7.9), *Xba*I (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), *Not*I (6 mM Tris-HCl, 150 mM NaCl, 6 mM $MgCl_2$, 1 mM DTT, pH 7.9).

Linear Dichroism. Linear dichroism (LD) was measured on a Biologic MOS-450 spectropolarimeter adapted for LD spectroscopy. The instrument was purged with nitrogen at a flow rate of 10 L/min. The bandwidth was 2 nm, and sample volumes were 50 μL for all measurements. The thermostated, microvolume Couette cell was manufactured by Crystal Precision Optics (Rugby, U.K.) and is available via Kromatek (Great Dunmow, U.K.). A schematic of the cell is shown in Figure 1. Reactions were initiated by mixing the enzyme

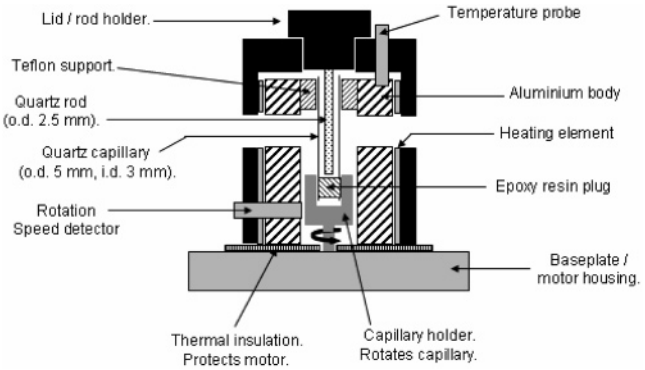


FIGURE 1: Schematic diagram of the microvolume, thermostated Couette cell (not to scale).

with the other reactants immediately prior to LD measurement. This results in a "dead time" of 10 s.

Wavelength scanning spectra were measured between 220 and 320 nm with a data interval of 1 nm and a data averaging time of 0.2 s at each data point; four spectra were averaged. Baseline spectra were subtracted from the raw data and were measured with the sample in the cell but with the rotation voltage set at 0 V (the capillaries are optically uniform and give the same baseline in all positions). Kinetic traces were measured at 260 nm, and the photomultiplier high-tension voltage (HT) was fixed at a value which gave an initial PMT output voltage of 5 V (dynamic range is 0−10 V). This was typically around 200 V, depending on sample concentration. The cell temperature was maintained at 37 ± 0.2 °C. Applied voltage to the motor was maintained at the stated value ± 0.02 V.

Data Analysis. The initial rates for the enzyme kinetics were taken from a linear least-squares fit to the linear part of the LD_{260} versus time curve, typically between 200 and 1000 s. This was after temperature and flow equilibration and when the kinetics had reached steady state but before the substrate becomes limiting. The rates were converted from LD units to concentration units using the slope of a calibration curve of LD_{260} versus concentration of linearized plasmid DNA. The rates were fitted to the Michaelis−Menten equation (eq 2) using Kaleidograph (Synergy Software, Reading, PA):

$$V = V_{max}[S]/(K_M + [S]) \quad (2)$$

where V is the velocity of the reaction, V_{max} is the maximum velocity, [S] is the substrate concentration, and K_M is the substrate concentration at which V is half-maximal ($V_{max}/2$).

RESULTS

The purified 6882 bp supercoiled DNA plasmid was first linearized using the restriction endonuclease *Eco*RI, which cuts at the sequence G∧AATTC (the caret indicates the position at which the phosphodiester bond is cleaved).

Signal versus Rotation and Concentration. The LD signal of linear DNA of 6882 bp varied with the speed of rotation of the quartz sample capillary. Spectra were measured over a range of voltages between 1 V (560 rpm) and 5 V (4800 rpm) with a 0.5 V increment (Figure 2A), and the LD signal at 260 nm was plotted against rotation speed (Figure 2B). The increase in the LD_{260} as a function of rotation speed is due to an increase in the alignment of the sample. It is clear

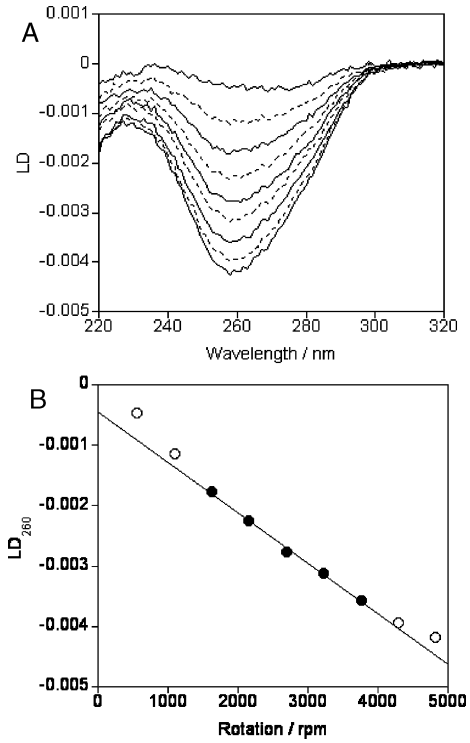


FIGURE 2: Linear dichroism of DNA at different rotation speeds. (A) LD spectra of 20 ng/μL DNA after linearization with *Eco*RI. The smallest signal is given by a rotation speed of 1 V (560 rpm), and spectra increase in magnitude with increasing voltage in 0.5 V increments up to 5 V (4800 rpm). (B) LD signal at 260 nm versus rotation voltage. The signal increases approximately linearly between 2 V (1600 rpm) and 4 V (3800 rpm).

that the signal approaches linearity with rotation speed inthe middle of the range but deviates from linearity at the extremes, as discussed below. The largest signal obtained within the linear range was at 4 V (3800 rpm), and for this reason subsequent data are all collected at this rotation speed.

The LD signal should also be directly proportional to concentration unless the viscosity of the solutions changes. To test this, we measured spectra of different concentrations of linear DNA at a constant rotation speed of 3800 rpm (Figure 3A). The LD_{260} signal was plotted against concentration of DNA bases, and a linear relationship was observed over the range tested (Figure 3B). The slope of the fit was −62.9/M of DNA bases (-8.66×10^5/M of linearized plasmid) with a good fit (R value of 0.99903). This was repeated for the supercoiled plasmid, and a value of −20.3/M of DNA bases (-2.79×10^5/M of supercoiled plasmid) was obtained. The difference between these values is −42.6/M of DNA bases (-5.87×10^5/M of plasmid that has been linearized), which provides a constant for conversion of the LD_{260} versus time plots that will be used for the kinetic analysis.

Kinetics of Restriction Endonuclease Reactions. The LD_{260} signal of DNA in the presence of the ten different restriction enzymes was measured over a period of 1 h. The LD signal as a function of time enables one to follow the effect of the restriction enzyme on the plasmid DNA. Three examples are shown in Figure 4. The enzyme *Eco*RI has a single recognition site in the plasmid and as such is expected to produce a linear DNA molecule. Figure 4A demonstrates that the observed kinetic trace shows an increase in negative LD_{260}. This is consistent with an increase in orientation with

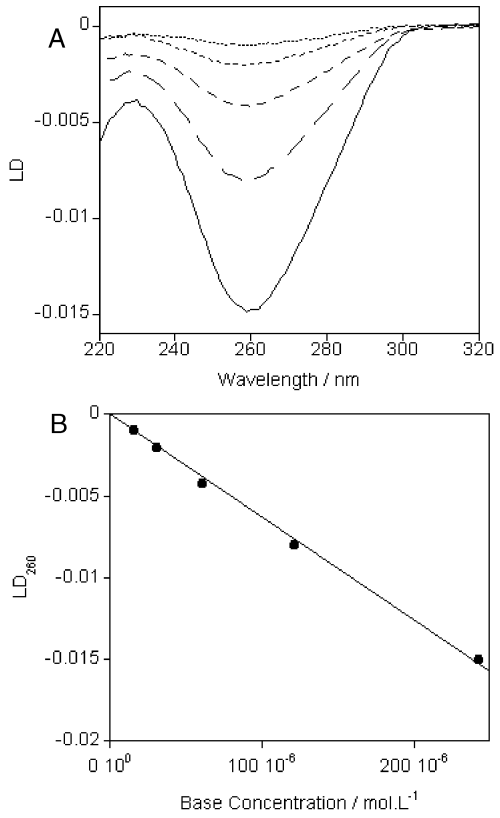


FIGURE 3: Linear dichroism of DNA at different concentrations. (A) LD spectra of linear DNA. Base concentrations were 15 μM (dotted line), 30 μM (short dashed line), 61 μM (dashed line), 121 μM (long dashed line), and 242 μM (solid line). (B) Data from (A) converted to the signal at 260 nm versus DNA concentration.

linearization of the plasmid, as predicted. This shows that LD can be used to follow restriction enzyme reactions in real time without the need for any dyes or labels.

The effect of enzymes with two cleavage sites in the plasmid DNA is also apparent in the LD traces of Figure 4B,C where the enzymes *Eag*I and *Bst*Z17I were used, respectively. In both cases the increase in the negative LD signal at earlier time shows the linearization of the plasmid. The decrease in the negative LD signal observed at later time reflects the second cleavage.

A further seven restriction enzymes were analyzed for their ability to cut the plasmid. Six of these enzymes had a single recognition site in the plasmid, and one (*Cla*I) had no site. We were initially surprised to find that the LD signal magnitudes at the end of some of the reactions were not as high as with *Eco*RI (Table 1), suggesting incomplete restriction of the sample. To verify by an established method the extent to which the enzymes cut, the samples were run on an agarose gel at the end of each LD kinetic run (Figure 5). It can be seen that *Sma*I (lane i) and *Xba*I (lane j) cut the DNA only partially. To a lesser extent this is also seen for *Nde*I (lane f). These gel results are thus entirely consistent with the LD results, and this confirms that the final LD signal can be used as a check for degree of digestion (most usually whether digestion is complete).

From the LD kinetic data, the time for half the substrate DNA to be cut ($t_{1/2}$) was calculated for each of the enzymes where this was appropriate, i.e., for those where efficient cutting at a single site occurred. The results are summarized in Table 1.

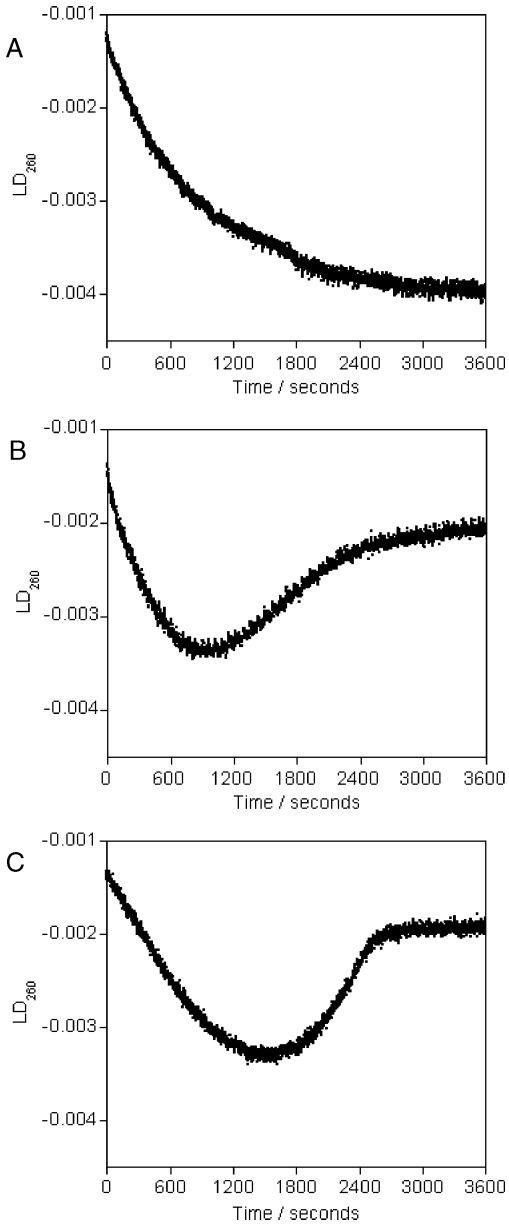


FIGURE 4: Kinetics of restriction digests of a circular, supercoiled plasmid. (A) Plasmid was digested with *Eco*RI, which has a single cut site in the DNA sequence. (B) Plasmid digestion with *Eag*I, which has two cut sites in the DNA sequence. (C) Plasmid digestion with *Bst*Z17I, which also has two cut sites in the DNA sequence.

Table 1: Restriction Digestion of Plasmid DNA by Restriction Enzymes with ≤1 Recognition Sites in the Plasmid DNA Sequence[a]

enzyme	final LD_{260}	cut	$t_{1/2}$/min
*Cla*I	−0.0010	none	N/A
*Eco*RI	−0.0040	complete	9.7
*Kpn*I	−0.0039	complete	0.83
*Nde*I	−0.0038	near complete	0.47
*Not*I	−0.0041	complete	1.2
*Nru*I	−0.0041	complete	0.83
*Sma*I	−0.0023	partial	N/A
*Xba*I	−0.0022	partial	N/A

[a] The enzymes were at a concentration of 0.4 unit·μL^{-1}, and the DNA was at a concentration of 40 ng·μL^{-1}.

To determine the kinetic parameters K_M and V_{max} using the initial rates, the enzyme concentration should be significantly below K_M. An *Eco*RI concentration of 0.2 unit·μL^{-1} was used for the kinetic experiments summarized in Figure

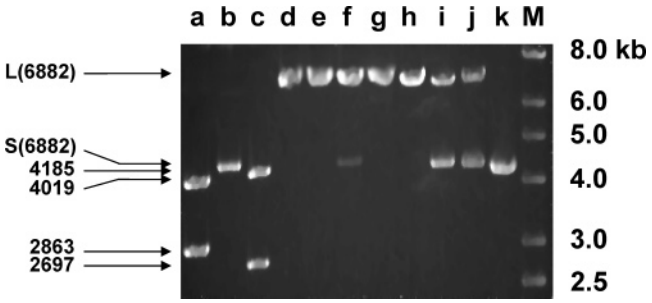


FIGURE 5: Agarose gel electrophoresis of restriction endonuclease reaction products: (a) *Bst*Z17I, (b) *Cla*I, (c) *Eag*I, (d) *Eco*RI, (e) *Kpn*I, (f) *Nde*I, (g) *Not*I, (h) *Nru*I, (i) *Sma*I, (j) *Xba*I, (k) uncut plasmid, and (M) molecular size markers. The arrows denote migration positions for the bands; linear is denoted by L and supercoiled by S.

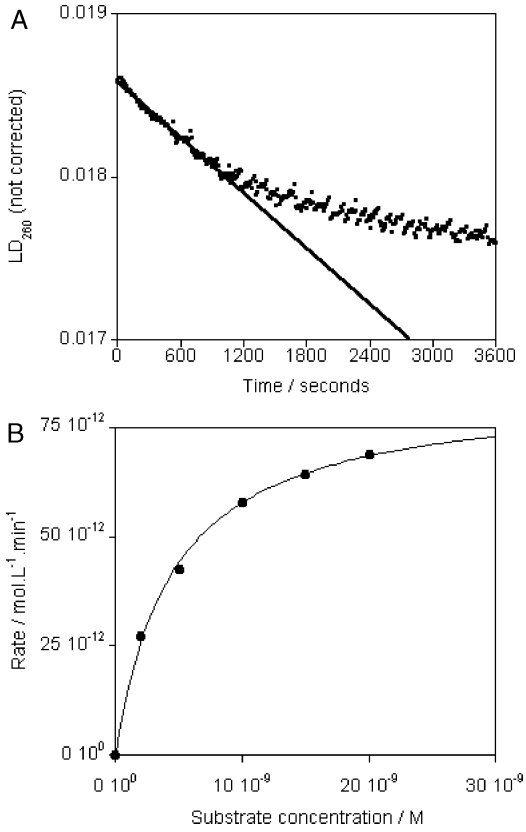


FIGURE 6: *Eco*RI restriction enzyme kinetics in detail. (A) The initial rate of product formation in the restriction digest was calculated from the slope of the steady-state part of the LD versus the time curve at different DNA (substrate) concentrations. This example is for 10 nM plasmid (138 μM bases). The rate was converted to concentration per unit time using the slope of the curve in Figure 3B. (B) Initial rate versus substrate concentration. The kinetic parameters K_M and V_{max} can be calculated by fitting this rectangular hyperbolic curve to eq 2.

6. Using a specific activity of 6×10^6 units·mg^{-1} (the average of the range provided by the manufacturers) and a dimer molecular mass of 62 kDa, this corresponds to an enzyme concentration of 538 pM. Substrate (plasmid) concentration was varied from 2 to 20 nM. The initial rates were measured as illustrated in Figure 6A, taking the slope of the LD_{260} versus time curve over the steady-state part of the reaction. These rates were converted to concentration units and were plotted versus substrate concentration. The resulting rectangular hyperbola is described by the Michaelis−Menten equation (eq 2). By fitting to the data using Kaleidograph (based on the Levenberg−Marquardt algorithm), the parameters V_{max} and K_M were found to be 84 pM·min^{-1} and 4.5 nM, respectively.

DISCUSSION

LD Signal versus Rotation and Concentration. Between 2 and 4 V the LD signal at 260 nm is linear with rotation voltage. The maximum sample alignment without inducing Taylor vortices and turbulent flow (*14*) is desired. At low rotation rates the rotation speed may not be stable, and at higher flow rates there may be turbulent flow. We therefore decided that the maximum voltage that should be applied must be within the linear response regime; this was 4 V in this case.

The LD signal of the linearized DNA plasmid of 6882 bp at constant rotation voltage varies linearly with concentration over at least the range 15−240 μM bases. This corresponds to approximately 1−20 nM plasmid. The linear relationship facilitates a conversion of LD units to plasmid concentration, so that enzymatic rates may be calculated in units of concentration per unit time by dividing by a constant: −42.6/M in this work. Note that the relationship of LD signal per unit concentration will depend on several factors, most notable DNA length, but also sequence (and hence flexibility), sample viscosity, and temperature; it should therefore be determined for the system being investigated.

Kinetics of DNA Cleavage by Enzymes with a Single Site in the Substrate Plasmid. We have demonstrated that one can follow the cleavage of plasmid DNA using LD spectroscopy to provide a real-time, continuous monitoring of the reaction. Some enzymes are known not to cut as efficiently as others. This is reflected in the magnitude of the LD signal at the end of the reaction (Table 1) and can be seen in Figure 6. Less efficient enzymes in this study include *Sma*I, *Xba*I, and to a minor extent *Nde*I. To confirm the LD signal correlated with cut efficiency, gels were run of the end point solutions.

To demonstrate that this method can be used quantitatively for determination of kinetic parameters, we undertook a more detailed study of the kinetics of one of the enzymes. We chose *Eco*RI because this is the most widely studied and used of the restriction enzymes, providing literature values for comparison. Our K_M value of 4.2 nM for *Eco*RI is in agreement with the literature values, which range between 1 and 10 nM for a range of plasmid substrates (*15*−*18*). The k_{cat} (or turnover number) equals V_{max}/enzyme concentration. The enzyme concentration was determined using a specific activity value of 6×10^6 units·mg^{-1}, which was the mean of the range ($2-10 \times 10^6$ units·mg^{-1}) given by the manufacturers (New England Biolabs U.K., Hitchin, U.K.). From our experiments k_{cat} therefore equals 0.05−0.26 min^{-1}. Note that the error in the measured value is due to uncertainty in the enzyme concentration and not error in the experimental data or fitting. This is at the lower end of reported k_{cat} values. However, the turnover number is variable and depends on the reaction conditions and will be affected by the amount of active enzyme present. The amount of active enzyme in a stock will also diminish over time. It is also possible using the assay presented here to detect incomplete digestion of the DNA. This is apparent from a smaller LD signal than for complete digestion.

Kinetics of DNA Cleavage by EagI and BstZ17I. Cleavage of the substrate DNA at two separate sites was also followed using LD. The increase in the negative LD_{260} signal that is observed in the first part of the curves in Figure 4B,C is due to the increased alignment that occurs upon linearization of the supercoiled plasmid. The subsequent decrease in the negative signal is due to the linear DNA molecule being cut a second time to give two shorter molecules. The shorter molecules align less well and thus give a smaller LD_{260}. The fragment sizes that result from DNA cleavage are 4185 bp + 2697 bp for *Eag*I and 4019 bp + 2863 bp for *Bst*Z17I. These fragment sizes are similar and should result in a similar LD_{260} upon complete digestion. This is observed at the end of the kinetic traces with values within 5% of each other. The time dependency of the LD for *Eag*I is much as one might expect, with an initial phase resembling the *Eco*RI curve and the gradually increasing effect of the second cleavage. The case for *Bst*Z17I, which is consistently reproducible, is unusual because the second cleavage event proceeds with a faster rate than the first but only after the completion of the first cleavage. This suggests that the second site may have a fast rate constant for cleavage but that the binding of the enzyme to this site is weak. This may be because of the site being buried and/or the sequence context. Some restriction enzymes require two DNA binding sites to initiate cleavage, resulting in complex kinetics; see, for example, the work on *Sfi*I from Halford's group (*19*, *20*). Work is in progress to mathematically model the more complex kinetic processes and to work out the mechanistic implications. The conclusion for the purposes of this work is that LD not only monitors the restriction reaction but also easily highlights any unusual kinetic aspects to the enzyme's activity.

CONCLUSIONS

We have demonstrated the versatility and ease of use of linear dichroism for the study of enzyme-catalyzed reactions of useful DNA substrates. There are many possible applications of this technique: not least, the study of restriction enzyme kinetics under different reaction conditions, batch to batch testing for restriction enzyme manufacturers, restriction enzyme-based assays, and following reactions in real time for molecular biology applications.

One of the key advantages of LD is that it is nondestructive, and the sample can be recovered after measurement. Digestion can be monitored in real time and the sample removed for subsequent genetic manipulation as soon as cleavage is complete. This has the advantage of saving time and avoiding "overdigestion", which may result in star activity where nonspecific sites are cut by some enzymes (*21*).

ACKNOWLEDGMENT

The authors thank Rhod and William Mortimore at Crystal Precision Optics (Rugby, U.K.) for the manufacture of linear dichroism cells. The plasmid was kindly supplied by Dr. Alex Connor from the University of Birmingham.

REFERENCES

1. Roberts, R. J., Belfort, M., Bestor, T., Bhagwat, A. S., Bickle, T. A., Bitinaite, J., Blumenthal, R. M., Degtyarev, S., Dryden, D. T., Dybvig, K., Firman, K., Gromova, E. S., Gumport, R. I., Halford, S. E., Hattman, S., Heitman, J., Hornby, D. P., Janulaitis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klaenhammer, T. R., Kobayashi, I., Kong, H., Kruger, D. H., Lacks, S., Marinus, M. G., Miyahara, M., Morgan, R. D., Murray, N. E., Nagaraja, V., Piekarowicz, A., Pingoud, A., Raleigh, E., Rao, D. N., Reich, N., Repin, V. E., Selker, E. U., Shaw, P. C., Stein, D. C., Stoddard, B. L., Szybalski, W., Trautner, T. A., Van Etten, J. L., Vitor, J. M., Wilson, G. G., and Xu, S. Y. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes, *Nucleic Acids Res. 31*, 1805−1812.
2. Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism, *Cell Mol. Life Sci. 62*, 685−707.
3. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
4. Ghosh, S. S., Eis, P. S., Blumeyer, K., Fearon, K., and Millar, D. P. (1994) Real time kinetics of restriction endonuclease cleavage monitored by fluorescence resonance energy transfer, *Nucleic Acids Res. 22*, 3155−3159.
5. Rarbach, M., Kettling, U., Koltermann, A., and Eigen, M. (2001) Dual-color fluorescence cross-correlation spectroscopy for monitoring the kinetics of enzyme-catalyzed reactions, *Methods 24*, 104−116.
6. Kettling, U., Koltermann, A., Schwille, P., and Eigen, M. (1998) Real-time enzyme kinetics monitored by dual-color fluorescence cross-correlation spectroscopy, *Proc. Natl. Acad. Sci. U.S.A. 95*, 1416−1420.
7. Hiller, D. A., Fogg, J. M., Martin, A. M., Beechem, J. M., Reich, N. O., and Perona, J. J. (2003) Simultaneous DNA binding and bending by EcoRV endonuclease observed by real-time fluorescence, *Biochemistry 42*, 14375−14385.
8. Zhang, Y. Y., Li, Q. G., Liang, J. X., and Zhu, Y. B. (2002) Hairpin probes for real-time assay of restriction endonucleases, *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai) 34*, 329−332.
9. Bier, F. F., Kleinjung, F., Schmidt, P. M., and Scheller, F. W. (2002) Determination of the turnover number of the restriction endonuclease EcoRI using evanescent wave technology, *Anal. Bioanal. Chem. 372*, 308−313.
10. Nilsson, P., Persson, B., Uhlen, M., and Nygren, P. A. (1995) Real-time monitoring of DNA manipulations using biosensor technology, *Anal. Biochem. 224*, 400−408.
11. Hillier, S. C., Frost, C. G., Jenkins, A. T., Braven, H. T., Keay, R. W., Flower, S. E., and Clarkson, J. M. (2004) An electrochemical study of enzymatic oligonucleotide digestion, *Bioelectrochemistry 63*, 307−310.
12. Waters, T. R., and Connolly, B. A. (1992) Continuous spectrophotometric assay for restriction endonucleases using synthetic oligodeoxynucleotides and based on the hyperchromic effect, *Anal. Biochem. 204*, 204−209.
13. Rodger, A., and Norden, B. (1997) *Circular Dichroism and Linear Dischroism*, Oxford University Press, Oxford.
14. Taylor, G. I. (1923) Stability of a viscous liquid contained within two rotating cylinders, *Philos. Trans. R. Soc. London, Ser. A 223*, 289−343.
15. Jack, W. E., Terry, B. J., and Modrich, P. (1982) Involvement of outside DNA sequences in the major kinetic path by which EcoRI endonuclease locates and leaves its recognition sequence, *Proc. Natl. Acad. Sci. U.S.A. 79*, 4010−4014.
16. Halford, S. E., and Johnson, N. P. (1981) The EcoRI restriction endonuclease, covalently closed DNA and ethidium bromide, *Biochem. J. 199*, 767−777.
17. Halford, S. E., Johnson, N. P., and Grinsted, J. (1980) The EcoRI restriction endonuclease with bacteriophage lambda DNA. Kinetic studies, *Biochem. J. 191*, 581−592.
18. Halford, S. E., Johnson, N. P., and Grinsted, J. (1979) The reactions of the EcoRI and other restriction endonucleases, *Biochem. J. 179*, 353−365.
19. Embleton, M. L., Williams, S. A., Watson, M. A., and Halford, S. E. (1999) Specificity from the synapsis of DNA elements by the SfiI endonuclease, *J. Mol. Biol. 289*, 785−797.
20. Wentzell, L. M., Nobbs, T. J., and Halford, S. E. (1995) The SfiI restriction endonuclease makes a four-strand DNA break at two copies of its recognition sequence, *J. Mol. Biol. 248*, 581−595.
21. Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M., and Boyer, H. W. (1975) Specificity of substrate recognition by the EcoRI restriction endonuclease, *Proc. Natl. Acad. Sci. U.S.A. 72*, 3310−3314.

BI0601712

CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200501012

An Estrogen–Platinum Terpyridine Conjugate: DNA and Protein Binding and Cellular Delivery

Michael J. Hannon,*[a] Philip S. Green,[a] Dianne M. Fisher,[b] Peter J. Derrick,[a] Jennifer L. Beck,[c] Stephen J. Watt,[c] Stephen F. Ralph,[c] Margaret M. Sheil,[c] Philip R. Barker,[a] Nathaniel W. Alcock,[a] Robert J. Price,[a] Karen J. Sanders,[a] Richard Pither,[d] Julie Davis,[d] and Alison Rodger*[a]

Abstract: A platinum metal complex in which terpyridine joins estradiol (via an ethynyl link) to a platinum with a labile ligand (chloride) has been designed, synthesised and its X-ray crystal structure determined. The aim of this work was to link a targeting motif (in this case estrogen) to a metal-based biomolecule recognition unit (the platinum moiety). The target molecule: 17α-[4′-ethynyl-2,2′:6′,2′-terpyridine]-17β-estradiol platinum(II) chloride (PtEEtpy) has been shown to bind to both human and bovine serum albumin (SA) and to DNA. FTICR mass spectrometry shows that the bimolecular units are in each case linked through coordination to the platinum with displacement of the chloride ligand. Circular dichroism indicates that a termolecular entity involving PtEEtpy, SA and DNA is formed. A range of electrospray mass spectrometry experiments showed that the PtEEtpy complex breaks and forms coordination bonds relatively easily. A whole cell estrogen receptor assay in an estrogen receptor positive cell (MCF-7) confirms binding of both EEtpy and PtEEtpy to the estrogen receptor in cells. The work demonstrates the concept of linking a targeting moiety (in this case estrogen) to a DNA binding agent.

Keywords: bioinorganic chemistry · DNA recognition · steroid conjugates · ternary complex · terpyridines

Introduction

Methods to take drugs to specific cells and to transport them into and localize them within those cells are an important goal in biomedical chemistry since they offer the potential to simultaneously reduce both drug dosage and side effects. Estrogenic steroids are attractive vectors for cellular targeting and delivery as it is known that they are transported through the blood (bound to steroid transport proteins) and across the cellular membrane[1] and that, by binding to the estrogen receptor (ER) in the cytoplasm, they are finally transported into the cell nucleus. Moreover, steroids localise in specific tissues (e.g. breast) which affords a level of targeting.[2] We have been interested in using estrogenic steroids to deliver metal centres into cells and report herein a steroidal metal complex in which an estrogen is used to deliver a platinum-based DNA-binding unit.

Nuclear DNA is believed to be the major biological target of many platinum-based agents,[3–5] generally due to their ability to undergo covalent coordination[6,7] or intercalative binding[8–14] with the DNA bases. Although the clinical anticancer agent cisplatin[5,15–20] is able to react with a wide range of biomolecules, its mechanism of action is largely attributed to the irreversible coordination of its platinum(II) centre (via the two chloro leaving groups) to the N^7 atoms of two guanine bases (and to a lesser extent, a guanine and an adenine base) of DNA, commonly in the form of 1,2-in-

[a] Prof. M. J. Hannon, Dr. P. S. Green, Prof. P. J. Derrick, Dr. P. R. Barker, Dr. N. W. Alcock, R. J. Price, Dr. K. J. Sanders, Prof. A. Rodger
Department of Chemistry, University of Warwick
Coventry, CV4 7AL (UK)
Fax: (+44) 24-7657-5795
E-mail: m.j.hannon@bham.ac.uk
a.rodger@warwick.ac.uk

[b] Dr. D. M. Fisher
School of Chemistry, University of Sydney
New South Wales, 2006 (Australia)

[c] Dr. J. L. Beck, S. J. Watt, S. F. Ralph, Prof. M. M. Sheil
Department of Chemistry, University of Wollongong
Wollongong, New South Wales, 2522 (Australia)

[d] Dr. R. Pither, J. Davis
Amersham Health, Amersham Place, Little Chalfont
Buckinghamshire, HP7 9NA (UK)

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

trastrand crosslinks.[5,16,21] Biologically active platinum(II) complexes which intercalate between DNA base pairs have also been developed.[22–25] Their intercalative ability results from an aromatic moiety of sufficient surface area either coordinated directly or linked indirectly to the platinum(II) centre which can π-stack between two adjacent DNA base pairs. Much of the early research concerning platinum(II) metallointercalators was conducted by Lippard and colleagues using complexes containing 2,2′:6′,2″-terpyridine (tpy).[13,26] Although the terpyridine-based platinum(II) complexes were shown to bind intercalatively to DNA, the complex $[PtCl(tpy)]^+$ was also observed to slowly form a metal coordination with DNA under certain ionic conditions. Replacement of the chloride by thiols[27] or an alkyl[28] gives inert complexes which exclusively intercalate, similarly replacement of the chloride with pyridine leads to a dicationic intercalating agent with higher DNA binding affinity[24] (although over time the pyridine can be displaced by nucleobases leading to the platination binding mode). Attempts to further improve the biological effectiveness of such terpyridine–platinum(II) complexes against tumour cells have led to the development of similar mono- and bifunctional compounds, including a number of drugs synthesized by Lowe et al.[29–32] containing 4′-substituted terpyridines and tpy-linked bispyridine ligands to create bis-intercalators with rigidly separated binding sites to explore DNA topology.[33] Pikramenou has recently linked two platinum–terpyridyl units to a dithiol substituted lanthanide complex to form very elegant visible and near-IR luminescent hairpin bis-intercalators[34] and the unit has also been attached to a thiol-substituted carborane.[35]

In addition to their DNA-binding, platinum(II)–terpyridyl $[Pt(tpy)X]^{n+}$ units are also known to bind to proteins, particularly at cysteine residues. They have been shown: to possess activity against ovarian carcinomas;[31] to possess antiprotozoanal activity[32] by being active against the parasites;[30,36] and to inhibit human thioredoxin reductase[29] and the renal sodium pump.[37]

Most platinum-based drugs are nonspecific in the types of cells they target. Thus, several attempts have been made to make them more specific for targeting tumour cells of particular organs.[5,15,38] One approach for treating hormone-dependent tumours is to link a cytotoxic species to a hormonal agent such as estrogen. As a result, the receptor of the natural hormone binds to the hormonal-like portion of the drug and selectively guides the normally non-selective cytotoxic segment of the drug to the hormone's target tissue, for example, the mammary tissue.[39–41] Early studies on estrogenic metal complexes focused primarily on cisplatin derivatives linked to estrogens through one of the two steroid hydroxy groups.[38,42–47] These two hydroxy groups are very important in receptor recognition[48] and, probably in consequence, there was little gain from such conjugation. Jaouen[49] explored conjugation of steroids to organometallic carbonyl compounds (for application in carbonylmetalloassays) and demonstrated that organometallics could be attached at the 17α-position of estradiol with retention of some steroid receptor binding ability. Katzenellenbogen explored attaching organometallics to the 7α-position with similar success.[50] Inspired by Jaouen's work we have explored the use of estradiol steroids as cellular and nuclear delivery vectors to deliver a variety of metal centres into cells. We have been able to demonstrate that, despite bearing cationic charge, 17α-functionalised steroidal metal complexes of a variety of metals (including palladium(II) and platinum(II)) are both delivered into the cell and retained some receptor binding (3–12% RBA).[51] Osella has more recently described an agent bearing a platinum–malonato unit attached through the 17α-position but which shows negligible receptor binding,[43] while Bérubé has prepared mixtures of estrogens with dichloroplatinum units attached at the 16α- and 16β-positions which show more promise.[52–54]

Lippard has demonstrated that steroid hormones can sensitise cancer cells to platinum drugs (through induction of HMG over-expression)[55] and has attempted to harness this affect through design of an estrogenic platinum(IV) agent intended to hydrolyse and reduce, thus forming cisplatin and an estrogen after entering the cell.[56]

The purpose of this work was to develop a molecule which combined: i) a steroid binding motif to facilitate cellular delivery to cells with estrogen receptors, and ii) a DNA binding motif designed for interaction with DNA. As steroids are transported through the blood by proteins such as human serum albumin, HSA, binding to HSA was also investigated. In this work we report the design, synthesis and biomolecule binding of 17α-[4′-ethynyl-2,2′:6′,2″-terpyridine]-17β-estradiol platinum(II) chloride (PtEEtpy), in which a terpyridine group links the steroid to the metal (Figure 1).

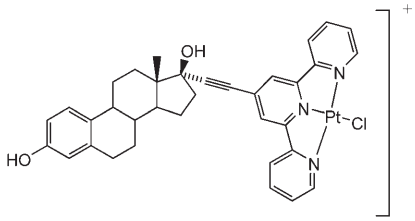


Figure 1. 17α-[4′-Ethynyl-2,2′:6′,2″- terpyridine]-17β-estradiol platinum(II) chloride, $[Pt(EEtpy)Cl]^+$.

Results and Discussion

Molecular design: Our design focused on attaching the DNA-binding unit to the 17α-position of the steroid. Previous work[51,66] has shown that this position locates the unit on the underside of the steroid away from the top recognition face and the hydroxy recognition units. Moreover, our previous work has shown that steroidal metal complexes conjugated through this point are transported into cells and retain some receptor binding.[66] Linking the steroid and metal complex via an alkyne linker is attractive due to its

synthetic feasibility and because, as a spacer unit, it introduces distance without introducing steric bulk. The platinum terpyridyl DNA-binding unit is attractive because its DNA binding is well studied and it has interesting biological activity. In addition, and as detailed below, Sonogashira couplings provide a powerful synthetic route to introduce an aromatic unit, making attachment of this unit to a steroid synthetically practical. We anticipated that the introduction of the large steroid group at the 4′-position of the terpyridine unit would prevent substantial intercalative insertion and thus drive the DNA-binding unit to adopt a covalent platination binding mode, thereby simplifying the binding.

Synthesis and characterization of [Pt(EEtpy)Cl]$^+$: The terpyridine derivatised estrogen ligand, EEtpy, was synthesized in good yields from ethynylestradiol and 4′-[{(trifluoromethyl)sulfonyl}oxy]-2,2′:6′,2″-terpyridine in a palladium-catalysed Sonogashira coupling. Purification (primarily removal of phosphine contaminants from the catalyst) was achieved by column chromatography on alumina. Mass spectral data reveal a parent ion EEtpy peak and NMR data are consistent with the coupled product. In particular, in the NMR spectrum the appropriate ratios of the integrals for the three aromatic estradiol protons with respect to the pyridyl protons are observed and the ethynyl proton (δ 2.8 ppm) from the starting steroid is absent. Recrystallisation of the ligand from a methanol/propan-2-ol mixture afforded crystals suitable for X-ray diffraction and the structure of this steroidal ligand has been determined.

The structure of the ligand (Figure 2) confirms that attachment of the terpyridyl unit through the 17α-position does indeed ensure the chelating unit lies below the steroid away from the two hydroxy groups and the upper face of the steroid which are important in receptor binding. The rings of the terpyridyl unit are approximately coplanar (torsion angles in the range 10–16°) and the nitrogens adopt the normal transoid arrangement observed in other crystal structures of free oligopyridines and which minimises lone pair-lone pair and dipole–dipole repulsions while maximising conjugation.[67–74] The planar terpyridyl ligand is oriented approximately parallel to the long axis of the steroid ligand; this orientation presumably permits the most efficient packing of the molecule. The presence of the terpyridyl unit does not interfere with the capacity of the steroid hydroxy groups to engage in hydrogen bonding with the phenolic and the tertiary hydroxy groups and methanol solvent molecules. Each steroid hydroxy group acts as an hydrogen-bond donor to one solvent molecule and an hydrogen-bond acceptor to another. The solvent-bridged hydrogen bonds link the steroids giving rise to steroid tube structures as shown in Figure 3. There are also a number of (very offset) π-stacking interactions and these and further packing representations are shown in the Supporting Information.

While coordinating a tridentate ligand to a square-planar metal might be expected to be facile, the synthesis of platinum(II) complexes of terpyridyl ligands can be challenging and indeed there have been studies devoted solely to this topic.[75,76] The synthesis is complicated by a tendency to form insoluble tetrachloroplatinate salts[75] and further complicated by different solid-state packing motifs that are believed to be responsible for dramatic differences in colour (yellow through red) and solubility. Moreover, no one synthetic route has been identified which is suitable for all differently substituted terpyridines. We first explored platinum cyclooctadiene dichloride as a starting material, following routes outlined by Annibale[75] and by Lowe.[76] Our first attempt afforded a clean pale yellow product of good solubility. However, we were unable to reproduce this synthesis, obtaining materials of varying colour and very poor solubility. Exploring variations on the approach, we were able to obtain complexes (evidenced by parent ion [Pt(EEtpy)Cl]$^+$ peaks in FAB MS) from both methanol or acetone solution and with and without pre-treatment with a silver salt to remove one or both chlorides. However, none of these routes gave products of good solubility (indeed most products obtained were insufficiently soluble for the NMR to be recorded). Finally we explored platinum bisbenzonitrile dichloride as a starting material. It was pre-treated with a silver(I) tetrafluoroborate or hexafluorophosphate salt to remove one of the coordinated chlorides. The reaction in dichloromethane led to a poorly soluble product, but by using

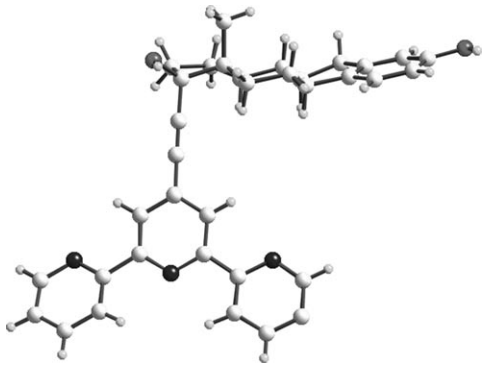


Figure 2. Crystal and molecular structure of EEtpy.

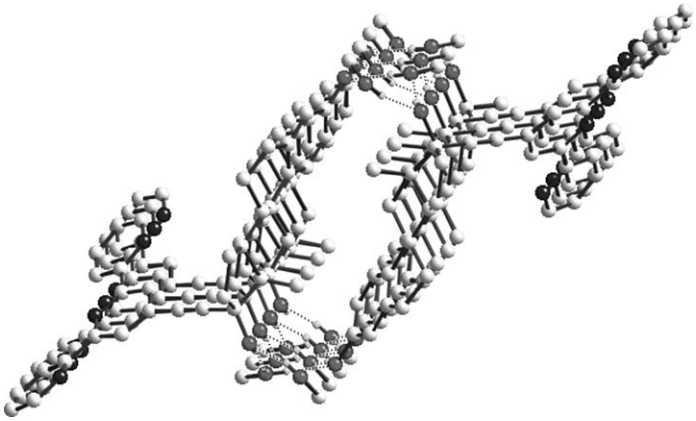


Figure 3. Hydrogen bonding of EEtpy with methanol solvent giving rise to steroid tube structures. For clarity, hydrogens not involved in the hydrogen bonds are omitted.

acetonitrile as reaction solvent a soluble orange tetrafluoroborate salt could be obtained. Anion metathesis was used to obtain the analogous chloride salt which has better solubility in alcohols and was used for the DNA binding studies. In their FAB and ESI mass spectra both the tetrafluoroborate and chloride salts show a peak at *m*/*z* 757 (with the correct isotopic distribution) corresponding to $[Pt(EEtpy)Cl]^+$. Partial microanalytical data are consistent with the formulations $[Pt(EEtpy)Cl]BF_4 \cdot H_2O$ and $[Pt(EEtpy)Cl]Cl \cdot 3H_2O$ and the high resolution FTICR mass spectrum (Figure 4) of the chloride salt confirms the identity of the cation with a monoisotopic mass of 756.1882 Da in agreement with the theoretical value of 756.1880 Da. The proton NMR spectra of the complexes show shifts in the pyridyl resonances consistent with coordination, and platinum satellites can be distinguished from the baseline for proton H_6. The UV/Visible spectrum of [Pt(EEtpy)Cl]Cl shows an absorbance centred around 415 nm ($\varepsilon = 2400\ mol^{-1}\,m^3\,cm^{-1}$) consistent with that for $[Pt(tpy)Cl]^+$ (400 nm), arising from a metal ligand charge transfer transition and again confirming platinum complexation.

Crystals of the tetrafluoroborate salt were grown from acetonitrile solution by the slow diffusion of benzene. The structure of the cation (Figure 5) reveals the anticipated *pseudo*-square-planar platinum(II) centre coordinated to a tridentate terpyridyl unit and a monodentate chloride. The terpyridine unit constrains the geometry at the metal and the Pt–N bond length to the central ring is consequently shorter than that to the outer rings (Table 1). The bond lengths and angles are similar to those observed in other platinum(II)–terpyridyl complexes.[77–80] The metal complex hangs below the estrogenic steroid, as in the free ligand, with the tpy unit again lying along the approximate direction of the long axis of the steroid. The complexes are stacked to form chains of platinum(II)–terpyridyl units, decorated with estrogenic steroids (Figure 6). There are no short Pt···Pt interactions (shortest contact > 4.3 Å) but rather the shortest contacts to the Pt centres are contacts to pyridyl rings above and below the square-plane of their coordination sphere. In each case this is to the C and N atoms of the central pyridyl ring, with the contact being shortest to the C (Pt···C 3.40 Å) for one such interaction and for the other to the centroid of the C–N bond (Pt···centroid 3.37 Å). All the pyridine rings are fully engaged in face to face π-stacking interactions and each tpy unit forms three stacking interactions to the tpy above and three to the tpy below (typical offset stacked interactions of coplanar rings with interplanar distances of the order of 3.4 Å and with centroid–centroid distances in the range 3.74–3.95 Å). This stacking motif is somewhat different from that usually observed: planar platinum–terpyridyl complexes do usually form extended stacks, however, within this stack it is usual to see platinum-platinum contacts, most usually Pt···Pt dimers[77–80] or polymers.[79] While previously reported Pt–terpyridyl stacks also contain some π-stacking interactions, not all rings have been fully engaged in that stacking. In this $[Pt(EEtpy)Cl]^+$ structure, the loss of potential Pt···Pt interactions (usually assumed to

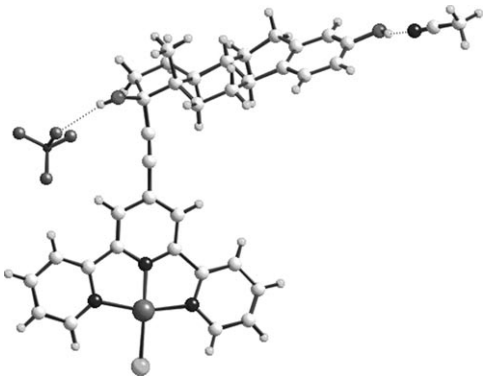


Figure 5. Crystal and molecular structure of $[Pt(EEtpy)Cl][BF_4]$.

Table 1. Selected bond lengths [Å] and angles [°] for $[Pt(EEtpy)Cl]BF_4$.

Pt(1)–N(1)	2.033(6)
Pt(1)–N(8)	1.901(6)
Pt(1)–N(14)	2.011(6)
Pt(1)–Cl(1)	2.295(2)
N(8)-Pt(1)-N(14)	81.1(2)
N(8)-Pt(1)-N(1)	81.1(2)
N(14)-Pt(1)-N(1)	162.1(2)
N(8)-Pt(1)-Cl(1)	178.6(3)
N(14)-Pt(1)-Cl(1)	99.78(19)
N(1)-Pt(1)-Cl(1)	98.06(19)

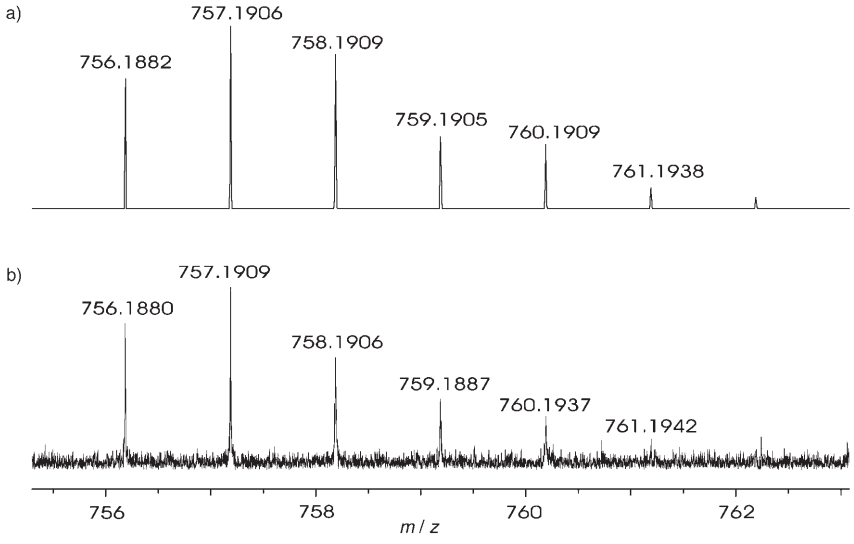


Figure 4. a) Theoretical and b) experimental FTICR mass spectrum of PtEEtpy chloride salt dissolved in water/acetonitrile/formic acid (49:49:2 v/v).

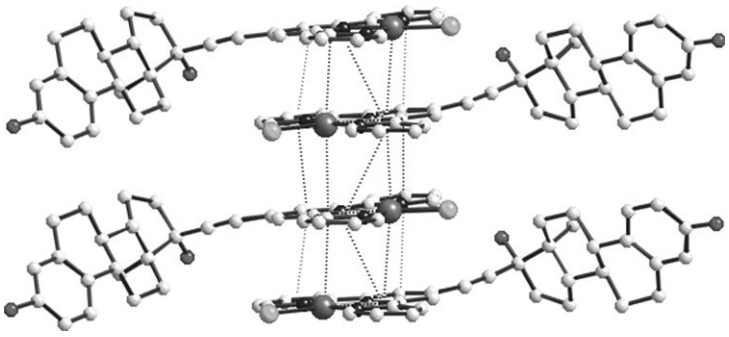


Figure 6. Chains of [Pt(EEtpy)Cl]+ cations formed through π–π and Pt-π interactions. Hydrogens are omitted for clarity.

be strong interactions)[79] appear to have been balanced by an increase in the π-stacking. This balance between multiple competing interactions is reminiscent of the "frustration" approach that we have recently explored as a means of supramolecular library design.[81,82] Indeed we have observed how such competition can lead to a variety of spiral or stacked structures in silver(I) terpyridyl compounds. While platinum–terpyridyl stacks are traditionally ascribed to Pt···Pt and π–π interactions, the contacts from the Pt centres to the rings above and below are particularly intriguing. They are very reminiscent of the η^2-type π-interactions typically seen for Ag···π interactions.[81,83–86] Moreover examination of previous platinum–terpyridyl structures reveals quite similar contacts at the points in the stacks where Pt···Pt contacts are absent. It is unclear whether these might represent additional weak bonding interactions or whether these contacts arise simply as a consequence of the aromatic stacking.

The steroids from adjacent chains are packed together to form an extended solid-state structure (see Supporting Information). As in the free ligand, both alcohol groups of the steroid engage in hydrogen bonding interactions in the crystal lattice, with the phenolic hydroxy hydrogen bonding with an acetonitrile solvent molecule (O···N 2.82 Å; OH···N 2.00 Å) and the tertiary alcohol with a tetrafluoroborate anion (O···F 2.93 Å; OH···F 2.15 Å). From the structure it is clear that the steroid would prevent intercalation into DNA via the central ring of the tpy unit, the initial binding mode observed for platinum complexes of unsubstituted tpy ligands and used so elegantly by Lowe[24,25,79] and by Pikramenou.[34] Some partial intercalation through a terminal pyridine ring might be possible but would not be extensive. By contrast the estrogen should have no steric effect on the alternative DNA-binding mode in which chloride is replaced by a DNA base (most usually N^7 of a guanine residue). The metal complex is buried beneath the steroid and does not interfere with the top face of the steroid nor with the hydroxy groups which, as the structure demonstrates, retain their ability to form hydrogen bonds.

Biomacromolecule binding: The molecular design concept of PtEEtpy was to use the steroid as a transport and localisation motif to deliver a DNA-binding platinum unit. Thus it was important to show that the complex bound to estrogen receptors in cells. The other key molecules of interest are human serum albumin (the main steroid transporter protein) and DNA.

Estrogen receptor whole cell binding assays: The relative binding affinities of the steroidal ligand and metal complex for ERs in viable MCF-7 cells were determined by a competitive radiometric binding assay. The assay involves competition of the steroidal conjugates with 16α-[^{125}I]-estradiol during incubation of viable cells in a medium containing a fixed concentration of 16α-[^{125}I]-estradiol and various concentrations of the competing steroidal ligand or metal complex. The bound 16α-[^{125}I]-estradiol is then extracted from the cells and measured using a gamma counter. From these data the concentration of competing chelate or metal complex required to displace half of the 16α-[^{125}I]-estradiol that would be bound to the ER in the cell in the absence of the test compound is calculated. This is denoted by the IC_{50} displacement measurement for the compound. Low IC_{50} values thus correlate with high binding affinities. This whole cell estrogen receptor assay depends on: i) the ability of the conjugate to be transported into the cell, and ii) the ability of the conjugate to bind to the estrogen receptor in the presence of other biomolecules in the cell which interferes with its binding to the ER (e.g. random metallation of protein amino-acid residues). In each experiment diethylstilbestrol was used as a control since it is known to have a high affinity for the ER (the concentration of DES at which 50% of the ^{125}I is displaced: IC_{50} = 0.6 nM in each experiment) and its use enables comparison of the receptor affinity to be made between assays.

The IC_{50} values (mean data from a minimum of two experiments) for the two molecules are: EEtpy: 80 nM and PtEEtpy: 500 nM. From these results it may be concluded that upon substituting the estrogen at the 17α-position with the tpy derivatives: receptor binding is retained (albeit at a lower relative binding affinity relative to estradiol, approximately 2% for the free ligand and 0.3% for the metal complex);[51] cellular delivery is achieved; and cationic charge (due to the Pt^{2+}) is not greatly detrimental to either delivery or binding. The magnitude of the receptor binding is weaker than the slightly more flexible complexes we have previously reported[51] where the platinum is chelated to a ligand with a single aromatic ring (to which the steroid is attached) and two flexible arms, as well as a single chloride ligand. The IC_{50} values for those complexes are ~15 nm. This may indicate that the planar terpyridine ligand is a little too large for the pocket adjacent to the receptor binding site though the assay does also probe other factors such as binding to other biomacromolecules and cellular transport. In

support of the latter option, a lysed cell-fraction ER assay gave a value of 6 nM for the IC_{50} of PtEEtpy (DES IC_{50} = 0.1 nM). The key conclusion is that receptor binding is retained upon conjugation of the steroid to the quite large terpyridine unit.

Spectroscopic studies of biomacromolecules binding: Circular dichroism (CD) is a measure of the chirality of a system and any changes in the observed signal when two molecules are mixed together is a sensitive probe of changes in the chiral environment of both chiral and achiral molecules. All the molecules of interest to this work are chiral, so, to use CD to probe change, it is essential that the intrinsic CD spectrum of each component (at the appropriate concentrations) is subtracted from the spectrum of the mixture. If the result is non-zero then there is an interaction. Linear dichroism (LD) is also a differential spectroscopy technique; in this case the light used is linearly polarized parallel and perpendicular to a sample orientation axis. In our experiments the DNA is oriented by flowing it (in a Couette cell).[62,63] DNA is the only molecule studied in this work which can be flow oriented, thus the LD signal is a useful indicator of whether anything is bound to the DNA, since for example, PtEEtpy will only give an LD signal if it is bound to the DNA. The magnitude of the DNA LD signal (when the ligands are not significantly contributing at 260 nm) is also a good indicator of whether the ligand bends[87] or stiffens the DNA.[63]

To probe whether PtEEtpy binds to either DNA or proteins the CD spectra of DNA, HSA, PtEEtpy in isolation and mixed together were measured. The ICD spectra of Figure 7 show that PtEEtpy binds both to DNA and to HSA. The perturbation to the PtEEtpy (P) spectrum (the region above 300 nm) by DNA (D) and HSA (H) are different (DP–D–P and HP–H–P in Figure 7) and that of the three-component mixture is distinct from that of any of the two-component mixtures or any simple combination of them (the simple average of the DNA–PtEEtpy and HSA–PtEEtpy ICD is illustrated in Figure 7). The ICD above 350 nm is largely (though not completely) due to the loss of the Pt–Cl ligand to metal charge transfer band (following loss of the Cl, see below). The terpyridine in-ligand band centred at ~325 nm is perturbed differently in all the mixtures. The spectra of the three-component mixtures are more similar, but not identical to an average of that induced by the HSA and DNA suggesting that the Cl is lost upon binding and the tpy is perturbed by both biomacromolecules. If the DNA is added first the ICD resembles that of the DNA–PtEEtpy ICD more closely than if the HSA is added first. This suggests that in each case there is a population of molecules adopting the preferred DNA binding and a population adopting the preferred HSA binding mode with the order of addition affecting this distribution. Thus it may be concluded that there is formation of some kind of heterobiomolecular complex of the three components: HSA–PtEEtpy–DNA. The complex thus formed is probably not a unique species.

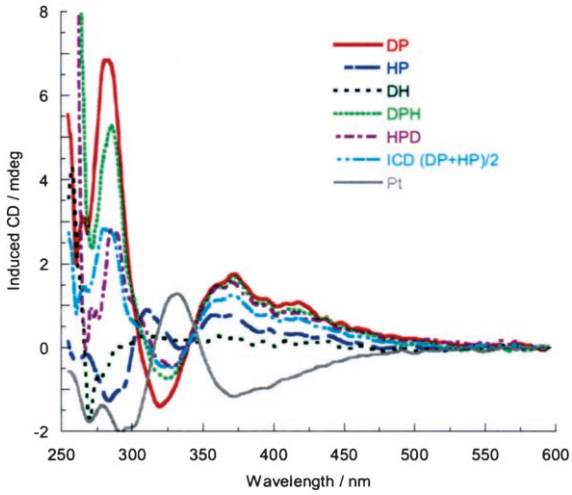


Figure 7. Induced CD spectrum for various mixtures of chiral components. In each case the spectrum is calculated by taking the baseline corrected spectrum of the mixture and subtracting the baseline corrected CD spectra of the relevant chiral components (PtEEtpy, HSA or DNA). DNA+PtEEtpy (denoted DP), HSA+PtEEtpy (denoted HP) and HSA+PtEEtpy+DNA (HPD). The notation denotes the order of mixing, e.g. DPH indicated the components were added in the order DNA then PtEEtpy then HSA. All spectra were collected in a 1 cm pathlength cuvette and data were smoothed over ±5 nm intervals. Concentrations are: PtEEtpy 30 μM, HSA was 16 μM, and ct-DNA 300 μM in 10 mM sodium cacodylate (pH 7) and 40 mM NaCl. The theoretical average of the DNA + PtEEtpy and HSA + PtEEtpy (denoted (DP+HP)/2) ICD spectra is also shown.

The LD spectra corresponding to the CD spectra of Figure 7 (see Figure S5, Supporting Information) confirm the simultaneous interaction of all three species. The LD of PtEEtpy+DNA shows a positive maximum at 470 nm and a negative maximum at 410 nm. Below 400 nm there are three negative bands with the signal at 258 nm being larger in magnitude than that for the DNA alone. The 330 nm tpy in-ligand band is small and negative which could be consistent with the tpy lying parallel to the DNA base pairs if the binding constant is relatively small—as would be expected either for intercalation or for covalent binding of the Pt to guanine N^7. The LD for the three component mixture (HSA + PtEEtpy + DNA) is smaller in magnitude than the PtEEtpy + DNA complex but similar in shape. This is further support for the formation of a heterobiomolecular array, and suggestive of the PtEEtpy adopting the same orientation on the DNA as in the absence of the HSA. The loss of LD magnitude in the ternary complex is due to a decrease in the orientation of the system, in contrast to the effect of HSA on DNA in the absence of PtEEtpy. The decrease could be caused by a "bending" or "kinking" of the DNA in the tri-molecular complex. The bending or kinking is indicative of a non-intercalative binding mode for PtEEtpy though it might be partially inserted into the edge of a base pocket.

ESI Mass spectrometry: The spectroscopic studies show both bi- and trimolecular complexation. Whether these com-

plexes are covalent or noncovalent cannot be determined from these experiments, though the data are suggestive of loss of the Cl ion from the Pt^{II} atom. If data can be collected, the high resolution of FTICR MS with a 9.4 T magnet provides the possibility of distinguishing covalent and noncovalent binding modes. Detailed experiments with a Qtof2 and non-self-complementary DNA sequences were also employed to further probe the complexation.

FTICR Mass spectrometry: BSA was used for these experiments as an HSA sample of sufficient purity to get good mass spectral data could not be obtained. The BSA spectrum of Figure 8 displays a charge state distribution ranging from $[M+28H]^{28+}$ to $[M+40H]^{40+}$. For each of the charge states shown, a degree of isotopic fine structure was observed, which allowed the average mass of BSA to be assigned as 66397 ± 1.4 Da, which is in good agreement with the theoretical average mass of 66397 Da calculated from the complete amino acid sequence.[88] Addition of PtEEtpy (30 μM) to BSA (30 μM) resulted in the formation of a BSA–PtEEtpy complex as shown in Figure 8 with a concomitant shift in the charge state distribution to lower charge states: the highest observed charge state in the un-complexed BSA was $[M+40H]^{40+}$ whereas the highest observed charge state in the complexed BSA was $[M+34H]^{34+}$. This shift in the charge state distribution upon the addition of PtEEtpy could reflect the BSA adopting either a more compact or a shielded conformation as a result of PtEEtpy binding. It was only possible to obtain the mass spectrum for the BSA–PtEEtpy complex in 8 K format as opposed to 512 K used for the free BSA and isotopic resolution for the complex was not achieved. However, the average mass of the observed complex, corresponded to BSA plus 721 Da, indicating that the complex contained only one PtEEtpy molecule without the chlorine attached. The theoretical ESI mass spectrum derived from the theoretical mass of BSA plus the mass of EEtpy minus chlorine was compared to the experimental spectrum and found to be in excellent agreement. Thus we conclude that the PtEEtpy has covalently bound to the BSA through the platinum rather than non-covalently through the estrogen as we had supposed (since BSA is a steroid transporter protein). This is not altogether surprising since the ratio of steroid binding sites to cysteine or methionine residues is ~1:15 for BSA.

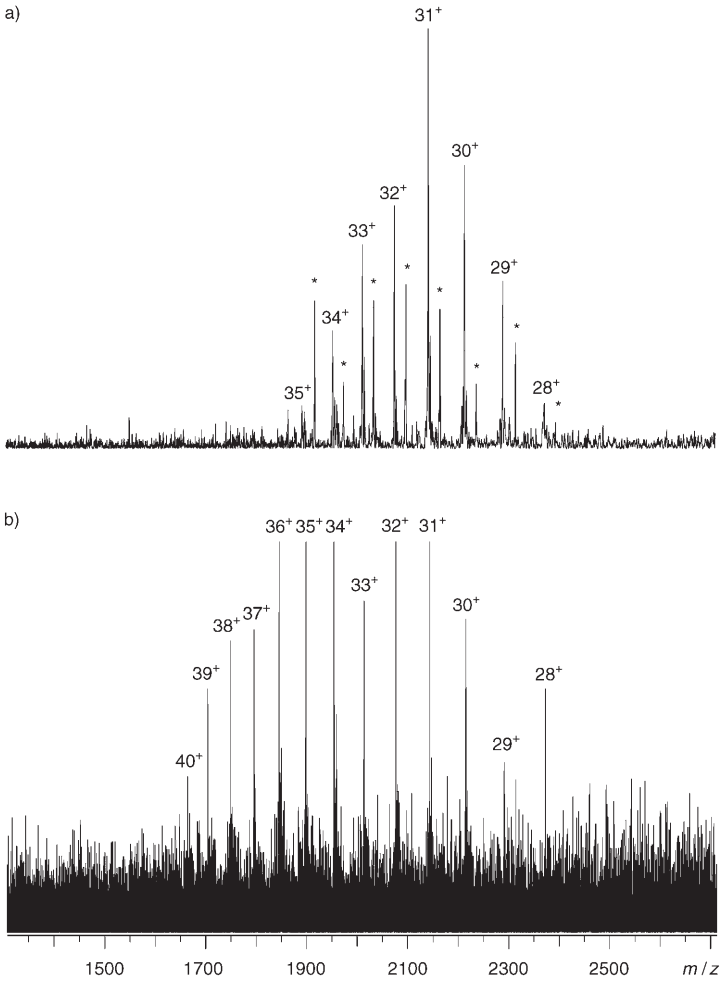


Figure 8. a) ESI mass spectrum of BSA (30 μM) + EPT (30 μM) and b) BSA (30 μM) in water/acetonitrile/formic acid (49:49:2 v/v/v). * Denotes peaks arising from EPT–BSA complex.

The proposed mode of action for PtEEtpy involves its ability to bind to DNA as well as serum albumin. The high resolution of the FTICR data means that ss and ds self-complementary oligonucleotides can be distinguished so the classical Dickerson–Drew dodecamer, (DD) 12-mer (5′-CGCGAATTCGCG-3′), was chosen to probe the DNA binding. The ion derived from this DNA with no sodium ions attached and without any protonation at the phosphodiester groups is denoted herein as $[DD]^{11-}$; the ion derived from the double-stranded Dickerson–Drew DNA (DD_2) would therefore, be assigned as $[DD_2]^{22-}$. DNA commonly retains sodium ions and frequently these sodium ions exchange with protons in solution resulting in various mixtures of cation combinations and charge states. In this work the concern was with maximizing the transfer of intact double stranded DNA from solution into the gas phase, so that the mode of PtEEtpy binding to DNA could be studied. The solvent adopted to maximize DD_2 and stability of the spray was water/acetonitrile/formic acid (49:49:2 v/v/v). CD was used to confirm that these conditions did not alter the DNA structure.

The negative ion mass spectrum of the DNA (200 μM oligonucleotide) obtained by nanospray ionisation shows (Figure 9) the presence of both DD and DD_2. There is a distribution envelope of sodium ions attached to the ss DD in the DD^{2-} and to the ds DD_2^{4-} charge states. Mass accuracies of less than 2 ppm were achieved for all the FTICR DNA spectra presented in this paper. Addition of PtEEtpy (50 μM) to the DNA (200 μM) resulted in several significant changes in the DNA spectrum compared to that of the free DNA (Figure 10). Firstly there was a shift in the

charge-state envelope for free DD_2 from low-charge states (3^-) to higher charge or more negative states (5^- and 6^-). Secondly, there was the appearance of peaks, which corresponded to a PtEEtpy–DD_2 complex in the M^{5-} and M^{6-} charge states. Thirdly, the single stranded DNA [DD], which was observed in the free DNA spectrum was no longer present. Finally, there was increase in the number of Na^+ ions attached to the free double stranded DNA (DD_2). Prior to the addition of PtEEtpy the charge state distribution for $[DD_2]^{4-}$ centred around the attachment of 12 Na^+ ions, upon the addition of PtEEtpy this increased to 16 Na^+ ions for free $[DD_2]^{4-}$. The poor signal intensities and broad isotopic distributions of the 5^- and 6^- PtEEtpy–DNA complex ions made assignment of the exact mass of the complex difficult despite repeated attempts. The average mass of the complex corresponded to DD_2 plus 721 mass units, indicating that the PtEEtpy+DNA complex contained only one PtEEtpy molecule without the chlorine attached, suggesting that the EEtpy covalently binds to the DNA through the platinum as in the case of BSA and as we anticipated in the molecular design. Thus we may conclude that it binds monofunctionally to the DNA, presumably to the N^7 of guanine,[12,13] that is, in the major groove.

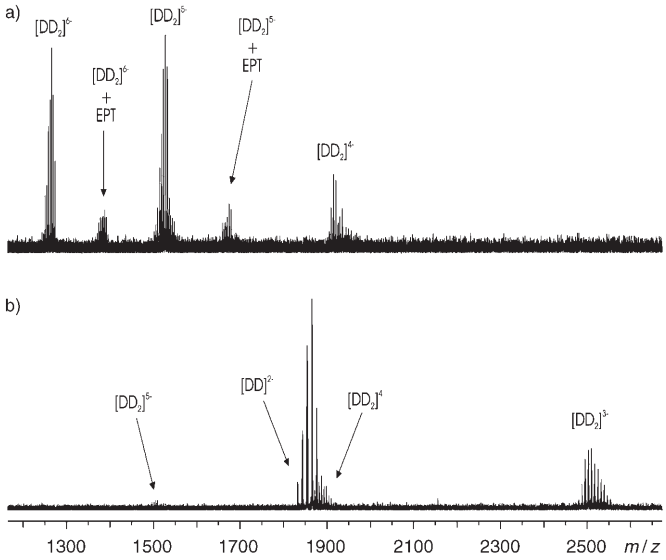


Figure 10. Negative ion nanoflow ionisation mass spectrum of a) Dickerson–Drew 12-mer (200 μM) + PtEEtpy (50 μM) in water:acetonitrile (1:1, v/v) and b) Dickerson–Drew 12-mer (200 μM). EPT denotes PtEEtpy.

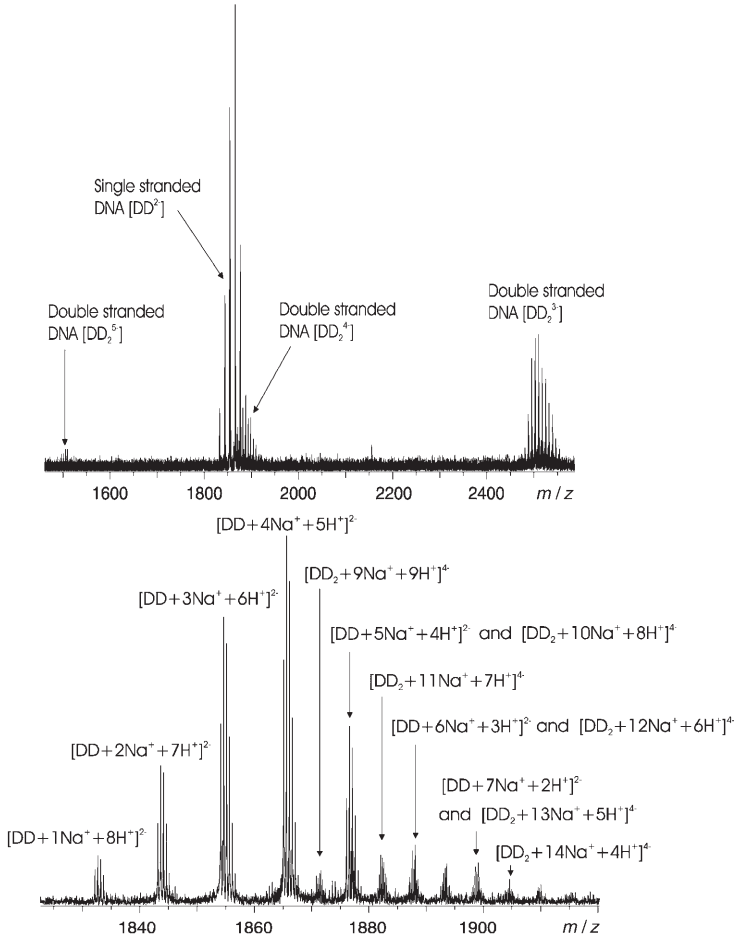


Figure 9. Top: Negative-ion nanospray ionisation FTICR mass spectrum of a Dickerson–Drew 12-mer (200 μM) in water/acetonitrile (1:3 v/v); bottom: expansion of the *m*/*z* 1830–1920 region of the spectrum. The Na^+ distribution centres around the attachment of 4 Na^+ for $[DD]^{2-}$ and 11 for $[DD]^{4-}$.

Binding to (5′-CCCATGCACACC-3′)·(5′-GGTGTGCATGGG-3′): The FTICR MS experiments produce data with isotopic resolution and enable identification of species present in solution, including differentiation of single- and double-stranded DNA and covalent or non-covalent binding. However, the experiments are non-trivial and attempts to observe the heterobiomolecular BSA–PtEEtpy–DNA complex in either positive ion or negative ion mode were unsuccessful. One of the issues appeared to be the small amounts of the PtEEtpy–DNA complex present in solution. Experiments were therefore undertaken to try to shift the reaction mixture towards the formation of the complexes following protocols developed for other coordinate bond forming DNA ligands.[65] A Qtof2 mass spectrometer was used to monitor the production of complexes. PtEEtpy was treated with SJW1 and the resulting complex was separated from unreacted ssDNA using HPLC.

Comparison of the −3 and −4 charge states of the SJW1–1PtEEtpy complex at *m*/*z* 1418 and 1063, respectively with −3 and −4 charge states of unreacted single stranded SJW1 at *m*/*z* 1177 and 889 (Figure 11) shows the effect of time and pH on the binding of PtEEtpy to SJW1. For example, at pH 5.0 (data not shown) the relative intensity of the 4^- complex ion becomes strongest after 6 h and after 122 h, the intensity of complex signals decreased significantly suggesting that the complex is not stable over long periods of time. At pH 7.0 the reaction is somewhat faster and the product more stable. At pH 8.5 the reaction is even faster though the amount of product is smaller. The pH 5 yields of the 1:1 complex are the greatest, but at this pH there are also noticeable amounts of 2:1 (*m*/*z* 1244 and 1659, 4^- and 3^- ions, respectively) and 3:1 PtEEtpy–DNA complex produced. Complexes of SJW1 with 2 PtEEtpy molecules (*m*/*z* 1244 and 1659, 4^- and 3^- ions, respectively) were also evident at the higher mixing ratio.

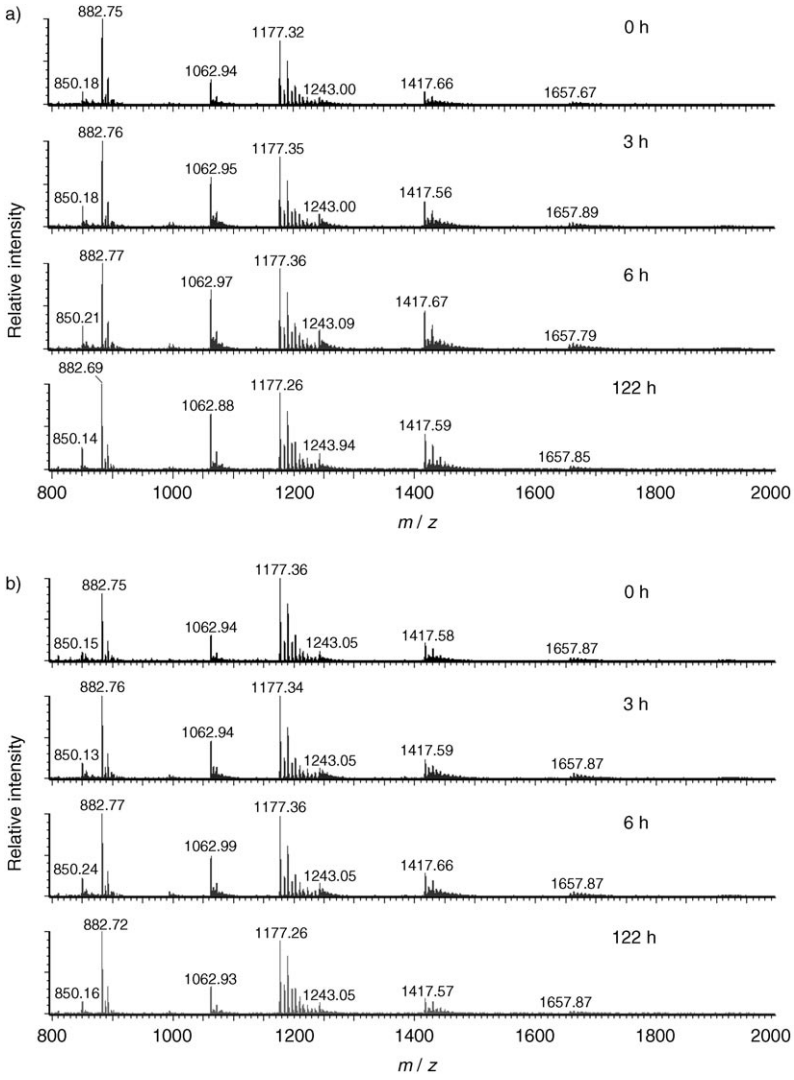


Figure 11. ESI-MS spectra acquired for SJW1–PtEEtpy at a ratio of 1:3 at a) pH 7.0 and b) 8.5 after various time increments as indicated in the figure.

The conditions for maximizing the 1:1 SJW1+PtEEtpy complex formation without significant amounts of 1:2 or 1:3 were therefore concluded to be: pH 7.0 with a ratio of 1:3 for 3 h. As judged by ESI-MS, these conditions seemed to produce an approximately equal amount of SJW1+1PtEEtpy complex and free SJW1 with small amounts SJW1+2PtEEtpy complexes. HPLC purification (Figure 12a) of the reaction mixture resulted in four major fractions (F1–F4). The mass spectrometry indicated fraction 1 comprised solely of the single-stranded SJW1, fraction 2 contained a mixture of SJW1 and the SJW1+PtEEtpy complex with an increased percentage of the complex being observed compared with the unpurified sample (Figure 12b). It is most likely that this fraction was eluted from the column as SJW1+PtEEtpy complex, however, by the time its mass spectrum was measured the solution had reached an equilibrium of bound and free ssDNA. Fraction 3 and 4 did not contain any DNA or complex (data not shown).

Fraction 2 (which contained ssDNA–PtEEtpy) was annealed with the complementary oligonucleotide SJW2. The ESI-MS spectrum of the resulting sample is shown in Figure 13. In this spectrum the high intensity of the 5^- ion of dsSJW1–SJW2 (*m*/*z* 1457) indicates it is present in very high amounts compared to the ds-SJW1–SJW2–PtEEtpy seen at *m*/*z* 1601. The relatively small amount of the dsDNA–PtEEtpy observed suggests that the annealing process (including the high temperature) has facilitated the dissociation of the SJW1–PtEEtpy complex.

Conclusion

We have synthesised a metallo–estrogen derivative that binds to serum albumin (a steroid transport protein), to DNA and to both protein and DNA simultaneously, showing that steroidal metal complexes can be used as bifunctional agents to assemble heterobiomolecular arrays. The X-ray diffraction crystal structure of PtEEtpy shows the shape of the molecule to be such that it is extremely unlikely to intercalate into the DNA but that the steroid would not prevent

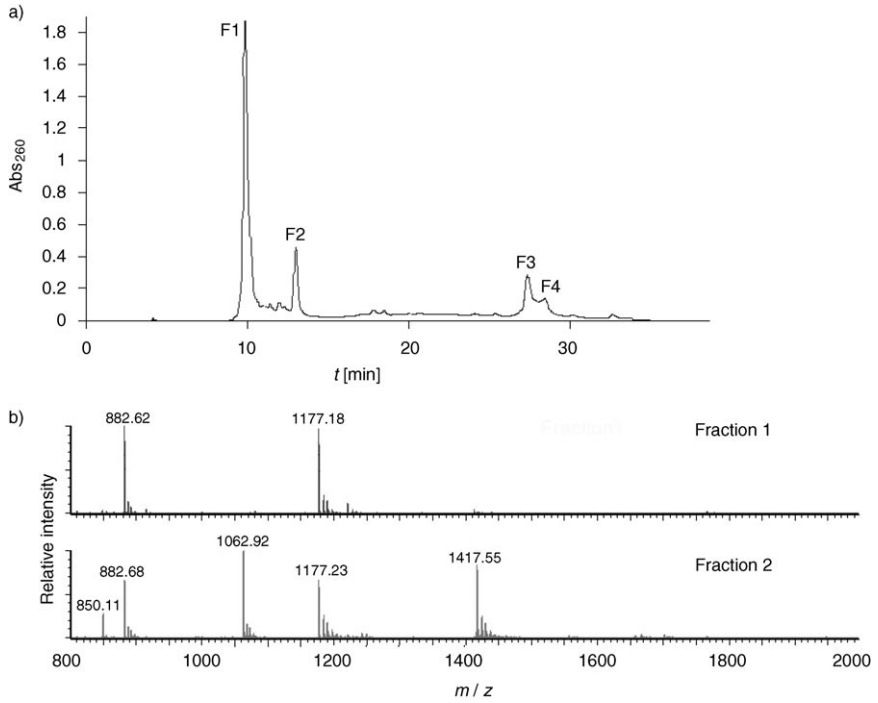


Figure 12. a) HPLC trace obtained from the purification of products from the reaction of SJW1 with PtEEtpy after reacting for 3 h at a ratio of 1:3 and pH 7.0; b) ESI-MS analysis of fractions 1 and 2.

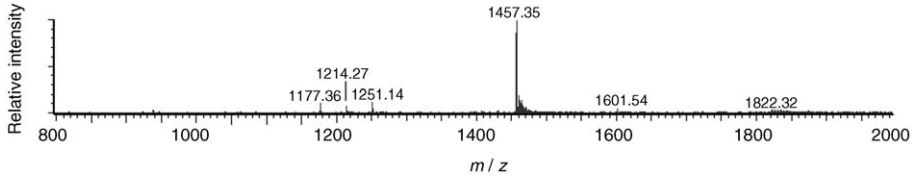


Figure 13. ESI mass spectrum of fraction 2 obtained from a HPLC separation of SJW1 reactions with PtEEtpy after annealing with SJW2.

it binding through metal coordination to the bases in the major groove of the DNA. The intermolecular interactions in the crystal are dominated by π–π stacking interactions and show no Pt–Pt interactions.

The FTICR mass spectra indicated that the bimolecular complexes of PtEEtpy with both DNA and BSA were metal coordination complexes formed by displacing the Cl^- ion from the metal. However, DNA and BSA cannot both bind in this manner in the trimolecular complex. The lability of the complex shown in the Qtof2 mass spectrometry experiments enables us to rationalize the fact that the CD spectroscopy showed a trimolecular array whose structure was in part dependent on order of addition of components: on the timescale of preparing and measuring the CD spectrum the complex could rearrange to a more stable form. The terpyridine is a very strong σ-donating π-accepting ligand so makes the position *trans* to it more labile than one might expect for a Pt–Cl bond. The shapes of the LD spectra suggest that the coordination bond to DNA is the one that is retained and that the protein effects a net bending of the DNA when it binds in or near the same site. Thus this monofunctional platinum agent with protein binding unit attached seems to have the ability to show a similar ultimate effect on DNA (kinking) as bifunctional cisplatin-like agents. The biological significance of the molecular design is shown by the in vitro whole-cell assay which showed that both the neutral metal-free ligand and the cationic platinum complex were transported and delivered to the cell, thus showing that the steroid is an effective delivery vehicle for cellular transport of metal centres. The work demonstrates that estrogenic steroids are effective delivery vehicles for DNA-binding/platinum agents and further investigations with other platinum units are ongoing.

Experimental Section

Synthesis

General: Starting materials were purchased from Aldrich, Avocado and Lancaster and used without further purification. ^{1}H NMR spectra were recorded on a Bruker DPX 300 MHz instrument. EI, CI and FAB mass spectrometry were carried out by the Warwick Mass Spectrometry Service using a matrix of 3-nitrobenzyl alcohol on Micromass Autospec spectrometers. ESI mass spectrometry was carried out by the EPSRC Mass Spectrometry Service Centre, Swansea on a Micromass Quatro II (low resolution triple quadrupole mass spectrometer. Infra Red spectra were recorded on a Bruker Vector 220 instrument fitted with ATR golden gate. UV/Visible spectra were recorded on a Jasco V-550 UV/Vis absorption spectrometer. Microanalyses were carried out by the Warwick Analytical Service on a Leeman labs CE44 CHN analyser. 4′-[{(Trifluoromethyl)sulfonyl}oxy]-2,2′:6′,2″-terpyridine was prepared in a three-step synthesis following the method of Potts.[57]

(17α-[4′-Ethynyl-2,2′:6′,2″-terpyridine])-17β-estradiol (EEtpy): 4′-[{(Trifluoromethyl)sulfonyl}oxy]-2,2′:6′,2″-terpyridine (33 mg, 8.66×10^{-4} mol), 17α-ethynyl estradiol (260 mg, 8.77×10^{-4} mol), palladium(II)bis(triphenylphosphine)dichloride (60 mg, 8.55×10^{-5} mol) and copper(I) iodide (9 mg, 4.73×10^{-5} mol) were mixed in degassed diisopropylamine (20 mL) under nitrogen, in the dark and stirred for 24 h. The solvent was removed in vacuo from the resulting orange solution and the solid was redissolved in diethyl ether (40 mL). The insoluble portion was removed by filtration and the filtrate removed to dryness yielding a light orange solid, which was dried over P_2O_5. Purification of the product was achieved by column chromatography using grade I alumina activated with 1% water and eluted with 1% methanol/chloroform giving a white solid (240 mg, 4.6×10^{-4} mol, 53%) which was dried over P_2O_5 under vacuum. Crystals suitable for X-ray diffraction were obtained by slow evaporation of an ethanol/propan-2-ol solution. ^{1}H NMR ($CDCl_3$): δ = 8.69 (br d, J = 4.7 Hz, 2H, H_6), 8.59 (br d, J = 7.9 Hz, 2H, H_3), 8.47 (s, 2H, $H_{3'}$), 7.86 (td, 2H,

$J=7.8$ Hz, H_4), 7.35 (ddd, 2H, $J=7.8$, 4.7 Hz, H_5), 7.14 (d, 1H, $J=8.3$ Hz, H_a), 6.63 (dd, 1H, $J=8.5$, 2.4 Hz, H_b), 6.54 (d, 1H, $J=1.9$ Hz, H_c), 2.79–1.23 (m, 21H, estradiol non-aromatic protons); IR: $\tilde{\nu}$ = 2930s, 2868s, 2812w, 1603w, 1584s, 1566m, 1547m, 1503m, 1469s, 1392s, 1286w, 1255m, 1119w, 1003w, 788s, 734m, 616w cm^{-1}; UV/Vis (CH_3OH): λ_{max} (ε): 317 (4000), 276 (14000), 248 nm (18000 $mol^{-1}\,m^3\,cm^{-1}$); EI+ MS: m/z: 527 $[M]^+$; CI+ MS: m/z: 528 $[M+H]^+$.

{(17α-[4′-Ethynyl-2,2′:6′,2″-terpyridine])-17β-estradiol}-chloroplatinum(II) tetrafluoroborate, [Pt(EEtpy)Cl]BF_4: Bis(benzonitrile)platinum(II) chloride (17 mg, 3.60×10^{-5} mol) in acetonitrile (5 mL) was added dropwise to a solution of silver tetrafluoroborate (9.7 mg, 5.0×10^{-5} mol) in acetonitrile (5 mL) and the resulting solution was heated at reflux for 16 h. Silver chloride, which had precipitated, was filtered off through Celite and to the filtrate was added EEtpy (19 mg, 3.61×10^{-5} mol). This solution was heated at reflux for a further 24 h after which the silver chloride which had precipitated was filtered off. The solution was reduced in volume by approximately two thirds and cooled in ice resulting in the precipitation of an orange solid (6.2 mg, 7.35×10^{-6} mol, 20%). Crystals suitable for X-ray crystallography were obtained by from acetonitrile solution by slow vapour diffusion of benzene. ^{1}H NMR (CD_3CN): δ = 8.89 (br d, $J=6.7$ Hz, H_6), 8.23 (t, $J=9.0$ Hz, H_4), 8.12 (br d, $J=7.4$ Hz, H_3), 8.10 (s, $H_{3'}$), 7.70 (br t, $J=7.4$ Hz, H_5), 7.00 (d, $J=7.9$ Hz, H_a), 6.42 (d, $J=10.0$ Hz, H_b), 6.37 (br s, H_c), 2.46–1.25 (m, 21H, estradiol non-aromatic protons); +ve FAB MS: m/z: 758 $[M-BF_4]^+$; elemental analysis calcd (%) for $Pt_1C_{35}H_{33}N_3O_2Cl_1B_1F_4{\cdot}1H_2O$: C 48.7, H 4.1, N 4.9; found: C 48.8, H 4.1, N 4.8.

{(17α-[4′-Ethynyl-2,2′:6′,2′-terpyridine])-17β-estradiol}-chloroplatinum(II) chloride, [Pt(EEtpy)Cl]Cl: Bis(benzonitrile)platinum(II) chloride (27 mg, 5.72×10^{-5} mol) in acetonitrile (5 mL) was added dropwise to a solution of silver hexafluorophosphate (14 mg, 5.54×10^{-5} mol) in acetonitrile (5 mL) and the resulting solution was heated at reflux for 16 h. Silver chloride, which had precipitated, was filtered off through Celite and to the filtrate was added EEtpy (31 mg, 5.89×10^{-5} mol). This solution was heated at reflux for a further 24 h after which additional silver chloride, which had precipitated out, was filtered off. Counter ion exchange was carried out by addition of *tert*-butyl ammonium chloride in acetonitrile (2 mL) resulting in the precipitation of an orange solid (31 mg, 3.96×10^{-5} mol, 69%), which was collected on a sinter and washed with cold acetonitrile (5 mL) and cold diethyl ether (5 mL). ^{1}H NMR (CD_3OD): δ = 9.10 (d, 2H, $J=6.8$ Hz, H_6), 8.55 (d, 2H, $J=7.8$ Hz, H_3), 8.52 (s, 2H, $H_{3'}$), 8.44 (t, 2H, $J=7.8$ Hz, H_4), 7.92 (v2H, $J=6.0$ Hz, H_5), 7.13 (d, 1H, $J=8.6$ Hz, H_a), 6.57 (d, 1H, $J=8.6$ Hz, H_b), 6.52 (s, 1H, H_c), 2.52–1.28 (m, 21H, estradiol non-aromatic protons); IR: $\tilde{\nu}$ = 3350s, 3067s, 2358w, 1713m, 1606s, 1556w, 1478w, 1418w, 1372w, 1292m, 1243m, 1097w, 1033m, 784m cm^{-1}; ESI-MS: m/z: 758 [Pt(EEtpy)Cl]$^+$; FAB+ MS: m/z: 758 [Pt(EEtpy)Cl]$^+$; UV/Vis (2% CH_3OH/H_2O): λ_{max} (ε) = 415 (2400), 334 (8000), 287 (13000), 259 (16000), 224 nm (18000 $mol^{-1}\,m^3\,cm^{-1}$); elemental analysis calcd (%) for $Pt_1C_{35}H_{33}N_3O_2Cl_2{\cdot}3H_2O$: C 49.6, H 4.6, N 5.0; found: C 49.5, H 4.5, N 4.9.

X-ray crystallography

Data were measured at 180 K with a Siemens-SMART-CCD diffractometer[58] equipped with an Oxford Cryosystem Cryostream Cooler.[59] Refinement used SHELXL96.[60]

Crystal data for EEtpy: $C_{37}H_{41}N_3O_4$, $M=591.73$, monoclinic, space group $C2$, $a=29.05630(10)$, $b=7.4338(2)$, $c=15.3049(3)$ Å, $\beta=99.128(2)°$, $V=3263.97(11)$ Å^3 (by least squares refinement on 3476 reflection positions), $T=180(2)$ K, $Z=4$, $\rho_{calcd}=1.204$ $Mg\,m^{-3}$, $F(000)=1264$; μ ($Mo_{K\alpha}$) = 0.078 mm^{-1}; crystal character: colourless plates; crystal dimensions $0.5\times0.4\times0.2$ mm; θ_{max} = 28.60°; *hkl* ranges were −38/30, −9/9, −5/20; 6189 reflections measured, 5559 unique [s(int) = 0.0194]; absorption correction: semi-empirical from equivalents; minimum and maximum transmission factors: 0.63, 1.0; no crystal decay.

Structure analysis and refinement: Systematic absences indicated space group $C2$, Cm or $C2m^{-1}$. The first was chosen because of the known chirality of the system and shown to be correct by successful refinement. The structure was solved by direct methods using SHELXS[61] with additional light atoms found by Fourier methods, including two lattice methanol molecules; their protons were located directly from difference maps, but were refined as restrained atoms. Hydrogen atoms were added at calculated positions and refined using a riding model with freely rotating methyl groups. Anisotropic displacement parameters were used for all non-H atoms; H-atoms were given isotropic displacement parameters equal to 1.2 (or 1.5 for methyl hydrogen atoms) times the equivalent isotropic displacement parameter of the atom to which the H-atom is attached. The absolute structure was determined from the known chirality of the molecule. Refinement of a delta-f" multiplier did not give any discrimination. Floating origin constraints were generated automatically. The weighting scheme was calculated $w=1/[\sigma^2(F_o^2)+(0.0551P)^2]$ where $P=(F_o^2+2F_c^2)/3$. Goodness-of-fit on F^2 was 0.842, $R1$[for 3703 reflections with $I>2\sigma(I)$] = 0.0411, $wR2=0.1009$; data/restraints/parameters 5559/1/404; largest difference Fourier peak and hole 0.113 and −0.145 e Å^{-3}.

Crystal data for [Pt(EEtpy)Cl]BF_4: $C_{43}H_{42}BClF_4N_4O_2Pt$, $M=964.16$, orthorhombic, space group $C222_1$, $a=6.7686(3)$, $b=25.2199(9)$, $c=47.0667(16)$ Å, $V=8034.4(5)$ Å^3 (by least squares refinement on 7971 reflection positions), $Z=8$, $\rho_{calcd}=1.594$ $Mg\,m^{-3}$, $F(000)=3840$; $\mu(Mo_{K\alpha})=3.620$ mm^{-1}; crystal character: red-brown plates; crystal dimensions $0.58\times0.18\times0.03$ mm; θ_{max} = 28.53°; *hkl* ranges were −5/9, −33/33, −60/55; 24648 reflections measured, 9361 unique [R(int) = 0.0802]; absorption correction: semi-empirical from equivalents; minimum and maximum transmission factors: 0.41, 0.93; no crystal decay.

Structure analysis and refinement: Systematic absences indicated space group $C222_1$. The structure was solved by direct methods using SHELXS[61] additional light atoms found by Fourier methods, including solvent benzene and acetonitrile molecules. Hydrogen atoms were added at calculated positions and refined using a riding model with freely rotating methyl and OH groups. Anisotropic displacement parameters were used for all non-H atoms; H-atoms were given isotropic displacement parameters equal to 1.2 (or 1.5 for methyl hydrogen atoms) times the equivalent isotropic displacement parameter of the atom to which the H-atom is attached. The known absolute structure was confirmed by refinement of a delta-f" multiplier. Absolute structure parameter $x=0.002(11)$. The weighting scheme was calculated $w=1/[\sigma^2(F_o^2)+(0.0400P)^2]$ where $P=(F_o^2+2F_c^2)/3$; goodness-of-fit on F^2 was 0.982, $R1$ [for 6395 reflections with $I>2\sigma(I)$] = 0.0634, $wR2=0.1105$; data/restraints/parameters 9361/0/504; largest difference Fourier peak and hole 2.883 and −3.759 e Å^{-3}; all the largest difference peaks are close to Pt; a number of relatively large discrepancies between observed and calculated structure factors are attributed to the length of the *c*-axis causing some overlap between peaks (but not enough to hinder structure solution or refinement).

Receptor binding and delivery

The binding to ERs in viable MCF-7 cells was investigated by a competitive radiometric binding assays against 16α-[^{125}I]-estradiol, denoted ^{125}I. The concentration of competing agent required to displace half the ^{125}I bound to the ER is presented as an IC_{50} displacement measurement. Low IC_{50} displacement values thus correlate to high binding affinities. EEtpy and PtEEtpy (10 mM) were dissolved in methanol to give a stock solution that was subsequently further diluted with assay buffer. Solutions were sonicated to ensure dissolution. A 2 mM stock solution of diethylstilbestrol (DES) was prepared in ethanol.

Cell growth conditions and receptor preparation: The estrogen receptor positive MCF-7 (human breast adenocarcinoma) cells (obtained from the American Type Culture Collection) were routinely grown at 37°C in 5% carbon dioxide/95% air and maintained by weekly passage in Minimum Essential Medium (MEM) (Gibco) supplemented with 2 mM L-glutamine, 50 IU per 50 μg mL^{-1} penicillin/streptomycin, 1% non-essential amino acids and 10% foetal bovine serum.

For radioligand binding assays MCF-7 cells were seeded in growth media at a concentration of 1×10^4 cells per well (1 mL) into 24 well culture plates. After a period of 24 h, media was removed from the wells and the cell monolayer was washed with PBS to remove residual media and serum. Estrogen free medium (1 mL) which consisted of MEM without phenol red (Gibco) supplemented with 2 mM L-glutamine, 50 IU per 50 μg mL^{-1} penicillin/streptomycin, 1% non-essential amino acids and controlled process serum replacement-1 (CPSR-1, Sigma) was added to

each well. Incubation was continued for a further three days after which time cells were used in radioligand binding studies.

Radioligand binding assays: Binding affinities were determined from a minimum of two independent assays in each case.

Viable cells assay: The medium was removed from a 24 well plate and monolayer washed with PBS (1 mL). The medium was replaced with 200 μL assay buffer (MEM without phenol red containing 0.1% BSA) containing 0.1 nM 16α-[^{125}I]-iodo-3,17β-estradiol plus or minus competing compound at desired concentration was added to the wells. The cells were incubated at 37 °C for 1 hour after which time assay reagents were removed and replaced with 500 μL equilibration buffer (5 mM sodium phosphate, 0.25 M sucrose, 0.5% BSA, 10% glycerol, pH 7.5). Cells were further incubated at 37 °C for 30 min. The buffer was removed and cells were washed rapidly with ice-cold equilibration buffer. Ethanol (1 mL) was added to each well to extract the radioactive component. After 30 min at room temperature, the ethanol was transferred to counting vials and radioactivity counted by a Wallac Wizard gamma counter.

Spectroscopy

Normal absorption spectra were collected on a Cary 1E or Jasco V-550 spectrometer in a 1 cm pathlength cuvette. Circular dichroism in a 1 cm or 1 mm pathlength cuvette and linear dichroism (LD), in a 1 mm path length Couette flow cell[62,63] spectra were collected on a Jasco J-715 spectropolarimeter adapted for LD measurements. All spectra were recorded at room temperature and were baseline corrected for any intrinsic signal in the cuvette or instrument. Induced CD (ICD) spectra were obtained by subtracting the baseline corrected spectrum of all chiral components of any mixture from the baseline corrected spectrum of the mixture of the components.

HSA was purchased from Sigma. Stock solutions (8 μM and 32 μM) were prepared by accurately weighing the appropriate mass of HSA (M_W 66437) and dissolving it in sodium cacodylate buffer (10 mM) and NaCl (40 mM), pH 7. These solutions were used without further purification. PtEEtpy stock solutions were prepared (by accurately weighing the solid) in 20:80 methanol (HPLC grade BDH):water (Elga Maxima 18.2 MΩ) v/v. DNA base molar concentrations were determined using $\varepsilon_{258\,nm} = 6600\ cm^{-1}m^3mol^{-1}$ for calf thymus DNA (highly lyophilized from Sigma).[64] The synthetic oligonucleotide (from Oswell, Southampton UK) and serum albumin (from Calbiochem, UK) concentrations were determined by weight assuming pure samples.

Mass spectrometry

Dickerson Drew DNA and serum albumin mass spectrometry experiments: Fourier transform ion cyclotron (FTICR) mass spectrometric experiments were undertaken with a 9.4 T Bruker Apex II instrument. Source parameters were carefully controlled during the experiments to minimise unwanted dissociation of the analyte and its complexes. Nanospray needles for the DNA experiments (Protana, USA) were coated in palladium and gold with a 50 μm internal diameter. The needles, which were tapered closed at the spray end, were opened by gently touching the needle against the end plate of the instrument. The exit of the tapered needle was measured under a microscope and found to be approximately 10–20 μm; a pressure of ~10 psi was applied to the proximal end of the nanospray needle. A capillary voltage of 300–500 V was typically used. Attempts to use HSA in the mass spectrometry experiments, even after desalting using a PD-10 column (Sephadex G25M, Amersham, Pharmacia Biotech) were unsuccessful. BSA was therefore used as it displays a very high sequence homology with HSA (76%), the same characteristic disulfide bond pattern and a very similar biological function. BSA and BSA–PtEEtpy 1:1 complexes were analysed by positive ion mode ESI. CO_2 was used as a drying gas in the source at a temperature of approximately 50 °C. The analyte was directly infused into the source at a flow rate of 1 $\mu L\,min^{-1}$. DNA and DNA:EPT (1:1) complexes were analysed using negative ion nano-flow ionisation. Due to its poor water solubility, a stock solution (200 μM) of PtEEtpy was prepared in 20% aqueous acetonitrile. For mass analysis the stock solution was diluted to 30 μM in water/acetonitrile/formic acid (49:49:2 v/v). For mass analysis the BSA was prepared to a concentration of 2 $mg\,mL^{-1}$ (30 μM) in water/acetonitrile/formic acid (49:49:2 v/v). A synthetic self-complementary Dickerson–Drew 12 mer (5′-CGCGAATTCGCG-3′) was purchased from Oswell (Southampton, UK) and used without any further purification. For mass analysis the DNA was prepared to a concentration of 200 μM in water/acetonitrile (1:3 v/v). The DNA–PtEEtpy complex was analysed at a ratio of four oligonucleotides per one PtEEtpy.

Preparation and ESI-MS analysis of ssDNA-PtEEtpy: A 1 mM stock solution of PtEEtpy was prepared in 100% methanol. Oligonucleotide 5′-CCCATGCACACC-3′ (SJW1) was dissolved in 10 mM NH_4OAc adjusted to pH 5.0, 7.0 or 8.5 with dilute acetic acid or ammonia solutions. Appropriate volumes of PtEEtpy stock solution were taken and mixed with single-stranded (ss) DNA (SJW1) giving ratios of PtEEtpy–ssDNA of 1:1, 3:1 and 6:1. The final concentration of ssDNA was 50 μM. Aliquots of these reaction mixtures were taken immediately and at 3, 6, 24 and 122 h after the time of mixing, diluted 50-fold in 50% CH_3CN in 5 mM NH_4OAc and placed on ice prior to electrospray ionization (ESI) mass spectrometry (MS) analysis. Samples were kept in the dark for all steps. Mass spectrometry of these samples was carried out using a Micromass Qtof2 mass spectrometer with a capillary voltage of 2600 V, cone voltage of 40 V, source block temperature 80 °C and desolvation temperature 120 °C. Typically 30 acquisitions were summed to obtain representative spectra.

Purification of ssDNA–PtEEtpy and annealing with complementary DNA: The double-stranded (ds) DNA–PtEEtpy samples were prepared as follows: PtEEtpy was mixed with SJW1 in 10 mM NH_4OAc pH of 7.0 at a ratio of 3:1. The final concentration of SJW1 was 0.4 mM, and the volume of the reaction mixture was 60 μL. After 3 h this solution was injected on to a Symmetry 300 C18 column (Waters) that had been equilibrated with 10 mM NH_4OAc. A linear gradient (0–65% CH_3CN in 10 mM NH_4OAc) over 30 min was used. Fractions were collected based on UV absorbance detection, dried using a vacuum centrifuge, and then redissolved in 0.1 M NH_4OAc, pH 8.5. Fractions were analysed by ESI-MS as above. This procedure was repeated a number of times to obtain concentrations of ssDNA-PtEEtpy suitable for annealing to SJW2 (5′-GGTGTGCATGGG-3′). An equimolar amount of SJW2 in 0.1 M NH_4OAc (pH 8.5) was then added to the SJW1 containing fractions. This produced ~60 μL of ~200 μM dsDNA. These solutions were held at 60 °C for ~15 min, and then allowed to cool slowly overnight. Samples were diluted in 0.1 M ammonium acetate, pH 8.5 prior to negative ion ESI-MS using a Micromass Qtof2 mass spectrometer as previously described.[65]

Acknowledgements

We thank the Swansea EPSRC National Mass Spectrometry Service for running electrospray mass spectra for compound characterization and EPSRC and Siemens for grants in support of the diffractometer. We are also grateful to the Engineering and Physical Sciences Research Council and Zeneca Agrochemicals for the award of studentships (to P.R.B. and P.G., respectively). MJH is the RSC Sir Edward Frankland Fellow (2004-5). James Lister Sinfield's help with data plotting is gratefully acknowledged.

[1] W. R. Miller, in *Estrogen and Breast Cancer*, Medical Intelligence Unit R. G. Landes Company and Chapman and Hall, **1996**.
[2] H. I. Jensen, H. I. Jacobson, *Biological Activities of Steroids in Relation to Cancer*, Academic Press, New York, **1960**.
[3] C. F. J. Barnard, *Platinum Met. Rev.* **1989**, *33*, 162–167.
[4] H. Baruah, C. G. Barry, U. Bierbach, *Curr. Top. Med. Chem.* **2004**, *4*, 1537–1549.
[5] M. A. Jakupec, M. Galanski, B. K. Keppler, *Rev. Physiol. Biochem. Pharmacol.* **2003**, *146*, 1–53.
[6] G. Admiraal, J. L. van der Veer, R. A. G. de Graff, J. H. J. den Hartog, J. Reedijk, *J. Am. Chem. Soc.* **1987**, *109*, 592–594.
[7] S. E. Sherman, D. Gibson, A. H. J. Wang, S. J. Lippard, *J. Am. Chem. Soc.* **1988**, *110*, 7368–7381.
[8] B. Lippert, *Cisplatin: Chemistry and the Biochemistry of a Leading Anticancer Drug*, Wiley-VCH, Weinheim, **1999**.

[9] S. J. Lippard, *Acc. Chem. Res.* **1978**, *11*, 211–217.
[10] Z. Guo, P. J. Sadler, *Adv. Inorg. Chem.* **2000**, *49*, 183–306.
[11] Z. Guo, P. J. Sadler, *Angew. Chem.* **1999**, *111*, 1610–1630; *Angew. Chem. Int. Ed.* **1999**, *38*, 1512–1531.
[12] M. Howe-Grant, S. J. Lippard, *Biochemistry* **1979**, *18*, 5762–5769.
[13] K. W. Jennette, S. J. Lippard, G. A. Vassiliades, W. R. Bauer, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 3839–3843.
[14] Z. Guo, P. J. Sadler, *Angew. Chem.* **1999**, *111*, 1610–1630; *Angew. Chem. Int. Ed.* **1999**, *38*, 1512–1531.
[15] T. Boulikas, M. Vougiouka, *Oncol. Rep.* **2003**, *10*, 1663–1682.
[16] M. A. Fuertes, J. Castilla, C. Alonso, J. M. Perez, *Curr. Med. Chem. Anti-Cancer Agents* **2002**, *2*, 539–551.
[17] G. Giaccone, *Drugs* **2000**, *59 (Suppl. 4)*, 9–17.
[18] I. Judson, L. R. Kelland, *Drugs* **2000**, *59 (Suppl. 4)*, 29–36.
[19] P. J. Loehrer, L. H. Einhorn, *Ann. Intern. Med.* **1984**, *100*, 704–713.
[20] J. Reedijk, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3611–3616.
[21] J. Reedijk, *Chem. Commun.* **1996**, 801–806.
[22] C. R. Brodie, J. G. Collins, J. R. Aldrich-Wright, *Dalton Trans.* **2004**, 1145–1152.
[23] H. L. Chan, D. L. Ma, M. Yang, C. M. Che, *ChemBioChem* **2003**, *4*, 62–68.
[24] A. McCoubrey, H. C. Latham, P. R. Cook, A. Rodger, G. Lowe, *FEBS Lett.* **1996**, *380*, 73–78.
[25] W. D. McFadyen, L. P. G. Wakelin, I. A. G. Roos, V. A. Leopold, *J. Med. Chem.* **1985**, *28*, 1113–1116.
[26] M. Howe-Grant, K. C. Wu, W. R. Bauer, S. J. Lippard, *Biochemistry* **1976**, *15*, 4339–4346.
[27] K. W. Jennette, J. T. Gill, J. A. Sadownick, S. J. Lippard, *J. Am. Chem. Soc.* **1976**, *98*, 6159–6168.
[28] G. E. Arena, L. M. Scolari, R. F. Pasternak, R. Romeo, *Inorg. Chem.* **1995**, *34*, 2994–3002.
[29] K. Becker, C. Herold-Mende, J. J. Park, G. Lowe, R. H. Schirmer, *J. Med. Chem.* **2001**, *44*, 2784–2792.
[30] S. Bonse, J. M. Richards, S. A. Ross, G. Lowe, R. L. Krauth-Siegel, *J. Med. Chem.* **2000**, *43*, 4812–4821.
[31] G. Lowe, A. S. Droz, T. Vilaivan, G. W. Weaver, J. J. Park, J. M. Pratt, L. Tweedale, L. R. Kelland, *J. Med. Chem.* **1999**, *42*, 3167–3174.
[32] G. Lowe, A. S. Droz, T. Vilaivan, G. W. Weaver, L. Tweedale, J. M. Pratt, P. Rock, V. Yardley, S. L. Croft, *J. Med. Chem.* **1999**, *42*, 999–1006.
[33] G. Lowe, A. S. Droz, J. J. Park, G. W. Weaver, *Bioorg. Chem.* **1999**, *27b*, 477–486.
[34] P. B. Glover, P. R. Ashton, L. J. Childs, A. Rodger, M. Kercher, R. M. Williams, L. De Cola, Z. Pikramenou, *J. Am. Chem. Soc.* **2003**, *125*, 9918–9919.
[35] J. A. Todd. , L. M. Rendina, *Inorg. Chem.* **2002**, *41*, 3331–3333.
[36] O. Inhoff, J. M. Richards, J. W. Briet, G. Lowe, *J. Med. Chem.* **2002**, *45*, 4524–4530.
[37] N. T. Ruddock, K. L. Arnett, B. J. Wilson, M. A. Milanick, *Am. J. Phys. Cell Physiol.* **2003**, *284*, c1584–c1592.
[38] M. P. Georaiadis, S. A. Haroutounian, K. P. Chondros, *Inorg. Chim. Acta* **1987**, *138*, 249–252.
[39] N. G. Knebel, E. v. Angerer, *J. Med. Chem.* **1991**, *34*, 2145–2152.
[40] J. P. DiZio, K. E. Carlson, C. J. Bannochie, M. J. Welch, E. v. Angerer, J. A. Katzenellenbogen, *J. Steroid Biochem. Mol. Biol.* **1992**, *42*, 363–373.
[41] C. Cassino, E. Gabano, M. Ravera, G. Cravotto, G. Palmisano, A. Vessieres, G. Jauoen, S. Mundwiler, R. Alberto, D. Osella, *Inorg. Chim. Acta* **2004**, *357*, 2157–2166.
[42] J. Altman, T. Castrillo, W. Beck, G. Bernhardt, H. Schonenberger, *Inorg. Chem.* **1991**, *30*, 4085–4088.
[43] H. Brunner, G. Sperl, *Monatsh. Chem.* **1993**, *83–102*, 83–102.
[44] E. M. Ehrenstorferschafers, N. Steiner, J. Altman, W. Z. Beck, *Z. Naturforschung Sect. B* **1990**, *45*, 817–827.
[45] C. Chesne, G. Leclercq, P. Pointeau, H. Patin, *Eur. J. Med. Chem.* **1986**, *21*, 321–387.
[46] O. Gandolfi, H. C. Apfelbaum, Y. Migrom, J. Blum, *Inorg. Chim. Acta* **1989**, *161*, 113–123.
[47] D. M. Spyriounis, V. J. Demopoulos, P. N. Kourounakis, D. Kouretas, A. Kortsaris, O. Antonoglou, *Eur. J. Med. Chem.* **1992**, *27*, 301–305.
[48] A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engeström, L. Öhmann, G. L. Greene, J.-A. Gustafsson, M. Carlquist, *Nature* **1997**, *389*, 753–758.
[49] H. E. Amouri, A. Vessieres, D. Vichard, S. Top, M. Gruselle, G. Jaouen, *J. Med. Chem.* **1992**, *35*, 3130–3135.
[50] M. B. Skaddam, F. R. Wüst, J. A. Katzenellenbogen, *J. Org. Chem.* **1999**, *64*, 8108–8121.
[51] A. Jackson, J. Davis, R. J. Pither, A. Rodger, M. J. Hannon, *Inorg. Chem.* **2001**, *40*, 3964–3973.
[52] V. Gagnon, M. E. St-Germain, C. Descoteaux, J. Provencher-Mandeville, C. Descoteaux, J. Provencher-Mandeville, I. Mathieu, V. Perron, S. Mandal, E. Asselin, G. Berube, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3927–3931.
[53] S. Parent, S. K. Mandal, E. Asselin, G. Berube, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5919–5924.
[54] V. Perron, D. Rabouin, E. Asselin, S. Parent, R. C. Gaudreault, G. Berube, *Bioorg. Chem.* **2005**, *33*, 1–15.
[55] Q. He, C. H. Liang, S. J. Lippard, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5768–5772.
[56] K. R. Barnes, A. Kutikov, S. Lippard, *J. Chem. Biol.* **2004**, *11*, 557–564.
[57] K. T. Potts, D. Konwar, *J. Org. Chem.* **1991**, *56*, 4815–4816.
[58] SMART user's manual, Siemens Industrial Automation Inc., Madison, WI, **1994**.
[59] J. Cosier, A. M. Glazer, *J. Appl. Crystallogr.* **1986**, *19*, 105–107.
[60] G. M. Sheldrick, SHELXL96, University of Göttingen (Germany), **1996**.
[61] G. M. Sheldrick, *Acta Crystallogr. Sect. A* **1990**, *46*, 467–473.
[62] A. Rodger, *Meth. Enzymol.* **1993**, *226*, 232–258.
[63] R. Marrington, T. R. Dafforn, D. J. Halsall, M. Hicks, A. Rodger, *The Analyst* **2005**, *130*, 1608–1616.
[64] R. D. Wells, J. E. Larson, R. C. Grant, B. E. Shortle, C. R. Cantor, *J. Mol. Biol.* **1970**, *54*, 465–497.
[65] R. Gupta, J. L. Beck, M. M. Sheil, *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2472–2480.
[66] C. Cassino, E. Gabano, M. Ravera, G. Cravotto, G. Palmisano, A. Vessieres, G. Jaouen, S. Mundwiler, R. Alberto, D. Osella, *Inorg. Chim. Acta* **2004**, *357*, 2157–2166.
[67] E. C. Constable, *Prog. Inorg. Chem.* **1994**, *42*, 67–137.
[68] M. H. Chisholm, J. C. Huffmann, I. P. Rothwell, P. G. Bradley, N. Kress, W. H. Woodruff, *J. Am. Chem. Soc.* **1981**, *103*, 4945–4947.
[69] K. Nakatsu, H. Yoshioka, M. Matsui, S. Koda, S. Ooi, *Acta Crystallogr. Sect. A* **1972**, *28*, S24.
[70] L. L. J. Merritt, E. D. Schroder, *Acta Crystallogr.* **1956**, *9*, 801–804.
[71] C. A. Bessel, R. F. See, D. L. Jameson, M. R. Churchill, K. J. Takeuchi, *J. Chem. Soc. Dalton Trans.* **1992**, 3223–3228.
[72] E. C. Constable, *Adv. Inorg. Chem. Radiochem.* **1986**, *30*, 69–121.
[73] E. C. Constable, S. M. Elder, J. Healy, M. D. Ward, *J. Chem. Soc. Dalton Trans.* **1990**, 1669–1674.
[74] E. C. Constable, A. M. W. C. Thompson, D. A. Tocher, M. A. M. Daniels, *New J. Chem.* **1992**, *16*, 855–867.
[75] G. Annibale, M. Brandolisio, B. Pitteri, *Polyhedron* **1995**, *14*, 451–453.
[76] G. Lowe, T. Vilaivan, *J. Chem. Res.* **1996**, 386–387.
[77] H. K. Yip, L. K. Cheng, K. K. Cheung, C. M. Che, *J. Chem. Soc. Dalton Trans.* **1993**, 2933.
[78] J. A. Bailey, M. G. Hill, R. E. Marsh, V. M. Miskowski, W. P. Schaefer, H. B. Gray, *Inorg. Chem.* **1995**, *34*, 4591–4599.
[79] J. S. Field, R. J. Haines, D. R. McMillin, G. C. Summerton, *J. Chem. Soc. Dalton Trans.* **2002**, 1369–1376.
[80] R. Büchner, C. T. Cunningham, J. S. Field, R. J. Haines, D. R. McMillin, G. C. Summerton, *J. Chem. Soc. Dalton Trans.* **1999**, 711–717.
[81] M. J. Hannon, C. L. Painting, E. A. Plummer, L. J. Childs, N. W. Alcock, *Chem. Eur. J.* **2002**, *8*, 2225–2238.
[82] F. Tuna, J. Hamblin, A. Jackson, G. Clarkson, N. W. Alcock, M. J. Hannon, *Dalton Trans.* **2003**, 2141–2148.

[83] M. Munakata, L. P. Wu, K. Sugimoto, T. Kuroda-Sowa, M. Maekawa, Y. Suenaga, N. Maeno, M. Fujita, *Inorg. Chem.* **1999**, *38*, 5674–5680.
[84] M. Munakata, L. P. Wu, L. Ning, T. Kuroda-Sowa, M. Maekawa, Y. Suenaga, N. Maeno, *J. Am. Chem. Soc.* **1999**, *121*, 4968–4976.
[85] M. Jansen, *Angew. Chem.* **1987**, *99*, 1136–1149; *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 1098–1100.
[86] M. Jansen, *Angew. Chem.* **1987**, *99*, 1136–1139.
[87] M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 879–884.
[88] K. Hirayama, S. Akashi, M. Furuya, K. Fukuhara, *Biochem. Biophys. Res. Commun.* **1990**, *173*, 693–646.

Received: August 19, 2005
Revised: March 10, 2006
Published online: July 31, 2006

Looking at long molecules in solution: what happens when they are subjected to Couette flow?†

Alison Rodger,[a] Rachel Marrington,[a] Michael A. Geeves,[b] Matthew Hicks,[ac] Lahari de Alwis,[a] David J. Halsall[d] and Timothy R. Dafforn[c]

Received 3rd April 2006, Accepted 24th May 2006
First published as an Advance Article on the web 14th June 2006
DOI: 10.1039/b604810m

Knowing the structure of a molecule is one of the keys to deducing its function in a biological system. However, many biomacromolecules are not amenable to structural characterisation by the powerful techniques often used namely NMR and X-ray diffraction because they are too large, or too flexible or simply refuse to crystallize. Long molecules such as DNA and fibrous proteins are two such classes of molecule. In this article the extent to which flow linear dichroism (LD) can be used to characterise the structure and function of such molecules is reviewed. Consideration is given to the issues of fluid dynamics and light scattering by such large molecules. A range of applications of LD are reviewed including (i) fibrous proteins with particular attention being given to actin; (ii) a far from comprehensive discussion of the use of LD for DNA and DNA–ligand systems; (iii) LD for the kinetics of restriction digestion of circular supercoiled DNA; and (iv) carbon nanotubes to illustrate that LD can be used on any long molecules with accessible absorption transitions.

Introduction

Many molecules in biological systems are long (*e.g.* DNA) or part of long assemblies of molecules (*e.g.* fibrous proteins or membrane-bound molecules). While it is clear that the structures of such moieties are important to their function, many of the powerful techniques of structural biology, including X-ray diffraction and NMR spectroscopy, are not well suited to high order macromolecular complexes. Microscopies of various kinds are extremely useful but if working on molecular scales, almost by definition, focus on single units (molecules or arrays of molecules) at any one time rather than the whole population and the samples are also not fully in the solution phase. The purpose of this article is to illustrate the advantages and disadvantages of solution-phase flow linear dichroism (LD) spectroscopy for structural characterization of solutions of long molecules ranging from biomacromolecular systems to carbon nanotubes. Our recent work in this area has been inspired by the increased availability of samples for structural analysis which in turn has led us to develop a range of new Couette flow cells to reduce both sample volume and light scattering and to enable experiments under tight temperature control.[1–5] The way has now opened to many more applications that were previously impossible due to restricted sample quantity, required conditions or molecule size.

[a] *Department of Chemistry, University of Warwick, Coventry, UK CV4 7AL*
[b] *Department of Biosciences, University of Kent, Canterbury, UK CT2 7NJ*
[c] *Department of Biosciences, University of Birmingham, Birmingham, UK B15 2TT*
[d] *Department of Clinical Biochemistry, Addenbrooke's Hospital, Cambridge, UK CB2 2QR*

† The HTML version of this article has been enhanced with colour images.

Much of the flow linear dichroism literature relates to nucleic acids. The DNA work published before the early 1990s has been superbly reviewed by Nordén, Kubista and Kuruscev in ref. 6. An earlier linear dichroism review by Nordén is more general in subject matter.[7] The aim of this article is to cover the developments that have happened in the last ten years or so and to give an overview of how linear dichroism can be used to study a wide range of long molecules from DNA to fibrous proteins to carbon nanotubes. The intent here is to be illustrative rather than comprehensive, though the references cited open up a literature trail that should lead to most available literature.

Linear dichroism

Linear dichroism (LD) is a spectroscopic technique that can be used with systems that are either intrinsically oriented, or can be oriented during an experiment by external forces.[6–8] LD measures the difference in absorption of light linearly polarised parallel and perpendicular to an orientation axis,

$$LD = A_{\parallel} - A_{\perp}. \quad (1)$$

In the case of flow orientation of long molecules, the orientation axis is the long axis of the molecule. If a transition moment (the direction of net electron displacement during an electronic transition) is aligned more parallel than perpendicular to the orientation axis, a positive LD signal is observed. Conversely, if the direction of electron movement is more perpendicular than parallel to the orientation axis, a negative LD signal is observed. When a transition moment is at an angle of 54.7° the LD signal equals zero, however effective the orientation. Some transition polarizations of relevance for biomacromolecules are shown in Fig. 1.[2,9–11]

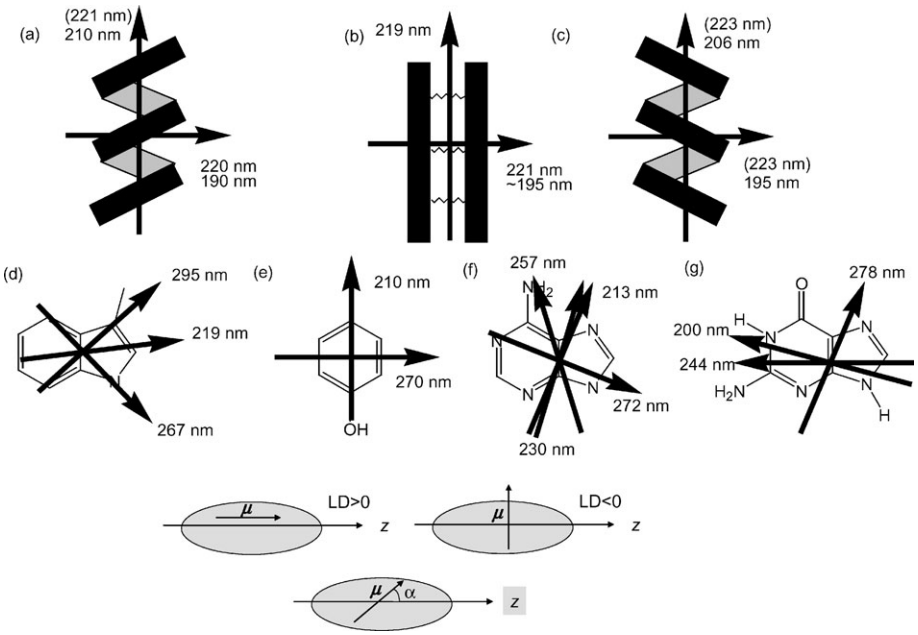


Fig. 1 The orientation of various polarization moments (a) α-helix (as determined by calculations[2]); (b) β-sheet (as determined by calculations[2]); (c) poly-proline type II helix (as determined by calculations[2]); (d) tryptophan;[9] (e) tyrosine; (f) adenine[10] and (g) guanine[11] chromophores together with LD schematic.

The reduced linear dichroism LD^r is often a convenient pathlength- and concentration-independent summary of LD data for quantitative analysis:

$$LD^r = \frac{LD}{A} = \frac{A_{||} - A_{\perp}}{A} = \frac{3}{2} S(3\cos^2\alpha - 1) \qquad (2)$$

where A is the absorption of the sample under isotropic conditions *i.e.* not oriented, and S a scaling factor (the orientation factor) defining the efficiency of the macroscopic orientation. For uniaxial rods, S equals 1 for perfect orientation, and 0 for random orientation. α (Fig. 1) is defined as the angle between the transition moment responsible for the absorption of light at a particular wavelength and the orientation axis. If either S or α is known the other can be calculated.

Molecules can be aligned using a number of techniques,[6–8,12] the most common being stretched film and flow orientation, which is the focus of this article. Flow orientation is ideal for biological molecules and others where the sample needs to be hydrated.[12–15] The orientation of long polymers is effected by the viscous drag created when a solution is flowed between narrow walls. A Couette flow LD cell, where a solution is placed in the annular gap between two cylindrical cells, one of which rotates and causes alignment of the molecules, has proved to be the most sample-efficient method of achieving this. A schematic diagram of a Couette cell is given in Fig. 2. For efficient alignment, the flow must be laminar, not turbulent.

Couette cells are derived from the work of Maurice Frédéric Alfred Couette and Henry Reginald Arnulph Mallock, who in the late 1800s independently developed a means to measure viscosity based on shearing a liquid between coaxial cylinders.[16–18,26] Since then there have been many applications of the Couette principle within the engineering and physics communities, the most notable example being Taylor–Couette flow which is a term used to define the flow between rotating cylinders.[27–29]

The modern incarnation of the Couette system developed in 1964 by Wada and Kozawa[20] involves one cylinder being inserted inside another and the solution flowing between the narrow gap (annular gap) as it is dragged by the rotation of one of the cylinders (Fig. 2). One of the cylinders must be transparent to the radiation being used and the other needs a transparent light path for the light which is incident radially on the cell (Fig. 2). The annular gap has typically been 500 μm, though experiments using a 50 μm annular gap have been reported.[13,15] In order to orient the sample it is necessary for the inner and outer cylinders to rotate at different speeds, therefore, creating a viscous drag and flow gradient in the solution; usually one cylinder is stationary. Previous studies have used either a rotating inner and fixed outer cylinder *e.g.* ref. 20 and 30, or a rotating outer and fixed inner cylinder *e.g.* ref. 31 and 32. The latter offers more flow stability but there are advantages for both options. In both cases it is necessary

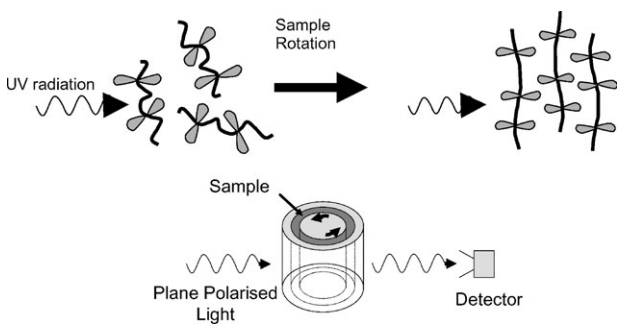


Fig. 2 Schematic diagram of a Couette flow cell.

that the light path through the cell allows optimal transmission, *i.e.* the rotating component and the region which is stationary where the light beam is incident need to be transparent to the radiation used. For example, for UV applications quartz is the most common material (though calcium fluoride has also been used[15]). The light beam can be either perpendicular[20] or parallel[31] to the axis of rotation. The perpendicular orientation is more common and is the one adopted in all studies detailed here.

Linear dichroism spectra are usually collected using a converted circular dichroism spectropolarimeter. The circularly polarised light of a CD spectropolarimeter is converted to linearly polarised light by the addition of a quarter wave plate, or as in our experiments, the photo elastic modulator 1/4 wave plate is converted to a 1/2 wave plate and the required alternating polarizations of light are produced directly by the instrument. Alternatively one can use polarisers in a normal absorption spectrometer and manually rotate the polarisers or the sample (though the latter is challenging for a Couette flow cell). The former option has significantly better signal to noise ratios. The linearly-polarised light passes through the sample and interacts with the oriented molecules to yield a net LD signal. As with all absorbance spectroscopy techniques a baseline must be subtracted from the sample spectrum. There are in principle two options for the baseline in a flow LD experiment. The first is to stop (or almost stop) the flow and measure the signal. With Couette flow cells this requires that either the rotating unit is optically uniform or that a very stable extremely low rotation speed is able to be maintained. The former is the case for our recently designed micro-volume Couette LD cells where the outer rotating cylinder is an extruded quartz capillary. Alternatively, one replaces the sample with buffer or solvent and rotates the unit as for the sample.

Fluid dynamics of flow in concentric cylinder flow cells

A fluid is a substance that will deform continuously when it is subjected to a tangential or shear force.[33] In the case of a Couette system, a velocity is imposed on the fluid contained between two concentric cylinders as one of these cylinders rotates (Fig. 2). The viscosity of a fluid (its resistance to shear or flow) will affect its dynamics. In all examples detailed in this work the solutions are aqueous so can be assumed to act predominantly like water, and therefore have constant viscosity independent of shear gradient, as discovered by Couette[26] (though high concentrations of sample increase the viscosity). Different flow states within the system exist, and it is these states that have been the subject of much study over the years since the work of Mallock and Couette.[16–19,26] The most notable work is that of Sir Geoffrey Ingram Taylor who investigated the flow stability of a viscous fluid when two cylinders rotate in the same direction and in opposite directions[27] and later went on to conclude that if only the outer cylinder rotates the flow is more stable than if only the inner cylinder rotates.[28] He also went on to study the effect of rotation speed and investigate the effect this has on fluid stability.

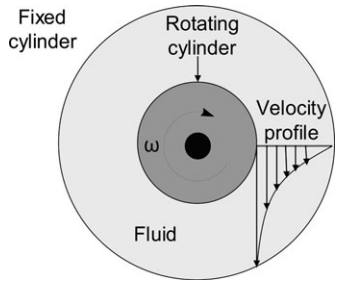


Fig. 3 Fluid flow profiles for inner rotating cylinder set-up, viewed from above. The longer the arrow describing the direction of the fluid flow, the faster the flow speed. Figure reproduced from R. Marrington, T. R. Dafforn, D. K. Halsall, M. Hicks and A. Rodger, *Analyst*, 2005, **130**, pp. 1608–1616—Reproduced by permission of the Royal Society of Chemistry.[4]

The spinning of either cylinder in a co-centric cylinder system, as in the LD cells of Fig. 2, causes the fluid in the annular gap to flow in a pattern dependent on many parameters in the system, including: the rotation speed of the cylinder, which cylinder rotates, the viscosity of the fluid, and the potential for interaction between the surface of the cylinder and the fluid or molecules in the fluid.[4,34] The flow profile for an inner rotating cylinder is given in Fig. 3.[33] Transitions between states are determined as functions of the inner and outer cylinder Reynolds numbers R_i and R_o, respectively. The Reynolds number is named after Osbourne Reynolds[35,36] and is a dimensionless variable that indicates the relative significance of the viscous effect compared to the inertia effect. The Reynolds number (Re) in the form applicable for a fluid in the annulus between two concentric cylinders is defined (kg s^{-1} m^{-1})

$$Re = \frac{\omega_1 r_1 \delta}{\nu} \tag{3}$$

where ω_1 is the angular velocity of one rotating cylinder in rad s^{-1}, r_1 is the radius of that rotating cylinder in m, δ is the gap width ($r_2 - r_1$) in m, and ν is the kinematic viscosity in m^2 s^{-1} (kinematic viscosity is equal to absolute or dynamic viscosity divided by density).[37] At high Re, the flow changes from laminar Couette flow to a series of more complex states, the first of which involves the appearance of Taylor vortices (vertically stacked toroids of flow around the inner cylinder) and the next involves wavy vortex flow which is exactly what the name says. These can be visualised by introducing light-reflecting suspended material into an LD cell as illustrated in Fig. 4. In practice we found that the Taylor vortices and wavy vortices could only be visualised in our original LD cell.[12] Our more recent and better engineered cells proved to have much more stable flow. Although Taylor vortex flow is still essentially horizontal as in Couette flow, the slow twisting rotation of the individual toroids is likely to have an effect on the alignment of molecules and it may depend on their flexibility. For example, we have observed that ethidium bromide bound to DNA is differently oriented in Couette flow and Taylor vortex flow regimes, presumably due to the different stiffness of the DNA with and without the intercalator ethidium bromide.[4]

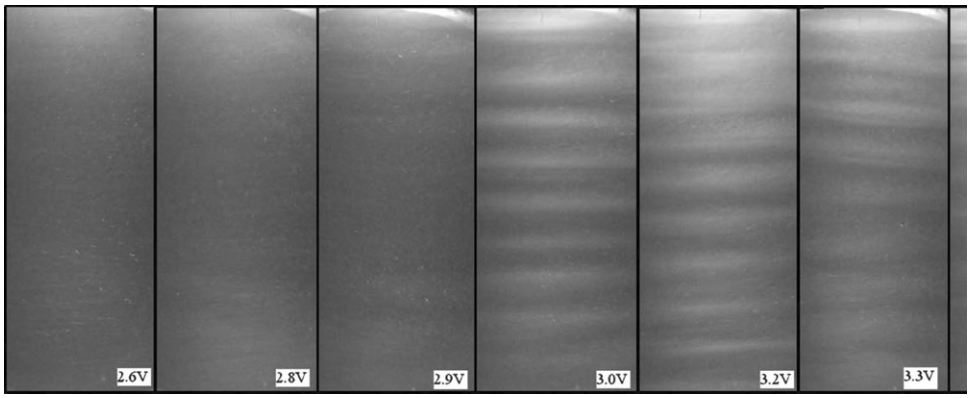


Fig. 4 Images of AQ-1000 rheoscopic fluid (titanium dioxide coated mica flakes in silicone oil, Kalliroscope gallery, 264 Main Street, P.O. Box 60, Groton, Massachusetts 01450, USA. http://www.kalliroscope.com) at a concentration of 5% in water in a concentric cylinder LD cell rotating at different speeds (speed increases linearly with voltage). A light source was shone onto the cell from above to illuminate the flakes. The AQ-1000 flakes align along the direction of fluid flow, providing an insight into the actual fluid flow in the cell. A 1 cm high slice through the middle of the cell window is illustrated at voltages from 2.6–3.5 V.

Light scattering correction

One of the challenges of working with high order macromolecular complexes, *e.g.* protein fibres or carbon nanotubes, is that they scatter light. The lenses of the microvolume capillary LD cells reduce this effect but it is still an issue, see *e.g.* ref. 38 and 39. The problem arises because the spectrometer simply measures the light that reaches the detector and assumes that which does not has been absorbed. Thus scattered photons are assumed to have been absorbed. A significant part of this problem can be removed by collecting the light as close to the LD cuvette as possible—this can be achieved by a number of methods including locating the photomultiplier tube (PMT) very close to the sample, increasing the angle of acceptance of the PMT, or locating a lens just after the sample to gather the light onto the PMT. However, even this does not always result in the desired flat line in the LD spectra outside absorbance regions. In such cases, the method of Nordh and others[39] can be used to subtract the scattering contribution to the signal as follows. Absorbance LD (LD^{A}) and background turbidity dichroism (LD^{τ}) contribute to the total LD signal as shown by eqn (5).

$$LD_{\text{total}} = LD^{A} + LD^{\tau}. \quad (5)$$

The wavelength dependence of LD^{τ} is accounted for by linear regression of eqn (5) to a function:

$$LD^{\tau}(\lambda) = a\lambda^{-k} \quad (6)$$

where k is a constant that has been shown generally to relate to the unpolarised turbidity (and usually takes values between 2.8 and 3.5), and a is a constant.[39] The implementation of this method is illustrated in Fig. 5.

Applications of linear dichroism

In the 1960s–1970s flow LD was recognised as a key technique and was used in the characterisation of fibrous proteins such as actin as well as for probing the orientation of DNA and bound ligands.[21–24] The limitations of instrument design and technology then, in particular the sample volume required (1–2 mL), and spectrometer design made it difficult to develop the technique further. The main applications reported in the literature lie in the analysis of nucleic acids,[25] mainly because of the research interests of groups with Couette flow cells and also because proteins can be difficult to orient (globular proteins), or generate artefacts due to light scattering caused by their large size. In the remainder of this article the applications of LD to a range of systems will be summarised. It should be noted that the discussion of DNA applications is far from complete due to space limitations.

Fibrous proteins

The cytoskeleton of a cell is a network of protein fibres whose repeat unit is a whole protein molecule and which form microtubules and microfilaments in the cytoplasm. The cytoskeleton is critical to cell motility (cell movement) and cell

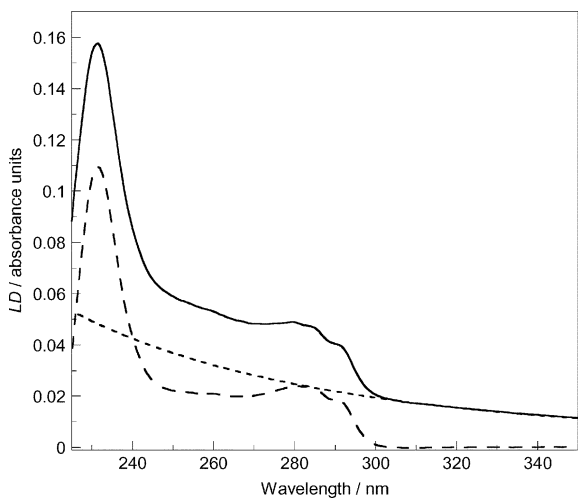


Fig. 5 A plot to show the method of light scattering correction when applied to a polymerised tubulin LD spectrum: the experimental data (—) (from which a baseline of the same sample in a non-rotating cell has been subtracted), the calculated turbidity LD, using a k value of 3.5, with a determined by rescaling the curve at 320 nm where there is no intrinsic absorbance (- - -), and the corrected data is (– – –).

morphology as well as processes such as cell division. It is not a rigid permanent structure, it is dynamic and constantly rearranges by polymerising and depolymerising to produce movement. There are three major types of protein fibre within the cytoskeleton of eukaryotic cells: actin filaments; intermediate filaments; and tubulin microtubules which are held together by protein–protein interactions. Apart from having structural roles within the cell, both tubulin microtubules and actin filaments also serve as 'tracks' for the biological movement of motor proteins—dyneins and kinesins along microtubules and myosins along actin filaments. More recently, interest has grown in fibrous proteins implicated in the mechanism of disease for such disorders as Alzheimer's and bovine spongiform encephalitis (prions). LD can be used to study these systems as discussed below. Actin was the first fibrous protein to be studied by Couette flow LD and recently we have improved the quality of data one can collect for such a system and also devised a range of new experiments. It is therefore an ideal case study to show possibilities of LD of fibrous proteins and is explored in some detail below. A brief review of work on other protein fibres is given after the discussion of actin.

Actin and actin-binding proteins. Actin exists as a globular monomer called G-actin which polymerises spontaneously to F-actin filaments in physiological conditions in the presence of ions such as Mg^{2+}, K^+ or Na^+. The process is reversible, in that when the ionic strength of the solution is lowered F-actin disassembles to G-actin monomers.[40] Actin binds 5′-adenosine triphosphate (ATP) or 5′-adenosine diphosphate (ADP)[41] which stabilise the fibre, but are not required for polymerisation. The three dimensional structure of actin molecules and actin filaments was solved in 1990 by Kenneth Holmes, Wolfgang Kabsch, and their colleagues.[42]

The polymerisation of actin has been a subject of interest for many years. The most common methods of analysis include viscometry, sedimentation, fluorescence spectroscopy, electron microscopy and light scattering, all of which have their merits. The initial application of Wada's Couette LD instrument in the 1960s and 1970s was to the study of F-actin solutions with a view to understanding the mechanism of muscle contraction.[21,24,43–45] Higashi and Oosawa, using the apparatus developed by Wada and Kozawa in 1964,[20] were able to investigate the orientation of aromatic amino acid residues (tyrosine and tryptophan) and bound adenine nucleotides. They reported that F-actin oriented by flow shows a negative dichroism at about 260 nm (due to adenine), a positive one at about 280 nm (tyrosine and tryptophan) and another negative one at about 295 nm (tryptophan).[21]

The use of LD to study actin did not proceed further at that time as it was concluded that the sensitivity of the apparatus needed to be improved to reduce light scattering and allow data collection at lower wavelengths.[45] Recent developments in spectrometer instrumentation and our new capillary Couette cells provide the required improvement and as shown below it is now possible to use LD as a technique either for, *e.g.*, the screening of actin–ligand binding or deciphering structural characteristics of the protein.[13]

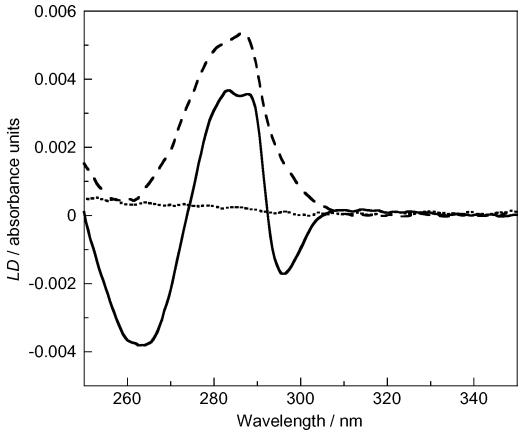


Fig. 6 Near UV LD spectra of actin (50 μM) (—); myosin S1 (50 μM) (- - -) and actin (50 μM)-myosin S1 (50 μM) (– – –) in KCl (15 mM); $MgCl_2$ (0.75 mM) and MOPS (3 mM) pH 7. The LD spectrum of unlabelled actin (50 μM)–myosin S1 (50 μM) has been corrected for baseline light scattering (k = 3.5, zeroing at 450 nm).

Fig. 6 shows a near UV LD spectrum of actin at 50 μM (see Appendix for experimental methods) which is consistent with those in the literature.[46] Far UV LD data for the backbone of actin was first collected using LD by Dafforn *et al.* in 2004.[13] The far UV LD spectrum of actin[46] at concentrations from 12–94 μM is shown in Fig. 7 and shows positive signals for both the n → π* transitions and the higher energy component of the π → π* transition, the lower energy component of the π → π* is evident as the 215 nm dip in the spectra. The LD therefore indicates that, on average, the α-helices in the fibre are oriented more perpendicular than parallel to the fibre axis in accord with literature data. Care must be taken in collecting low wavelength data for a system such as actin where light scattering is significant as illustrated in Fig. 7 where diluting the sample leads to an apparent wavelength shift of the maximum LD signal of ~30 nm to shorter wavelengths. When the dilution effect is simply to reduce the signal intensity according to the Beer Lambert law, then we can conclude the spectrum is real. From Fig. 8 it can be concluded that the LD wavelength maximum of the higher energy component of the π → π* transition is at 197.5 nm, in contrast to the previously reported maximum at 200 nm.[13] In all such experiments on fibrous proteins in general and actin in particular it is important to investigate the stability of the formed fibres to sheer forces by measuring the LD as a function of rotation speed.

The monomeric G-actin has no orientation in the flow, therefore the LD signal (shown in Fig. 7b at 222 nm) provides a convenient signal to follow the process of polymerization.[13] LD also provides a way to monitor the binding of globular proteins to fibrous proteins since globular molecules will not normally orientate in a flow cell, but upon binding to actin they do. The myosin motor domain (or subfragment 1, S1) binds tightly and stoichiometrically to F-actin. On its own, it has no LD spectrum (Fig. 6) but when bound to actin its near UV spectrum is distinct from the F-actin spectrum with loss of the negative peaks at 295 and 275 nm and enhanced 285 and 290 nm positive peaks reflecting the differences in the average

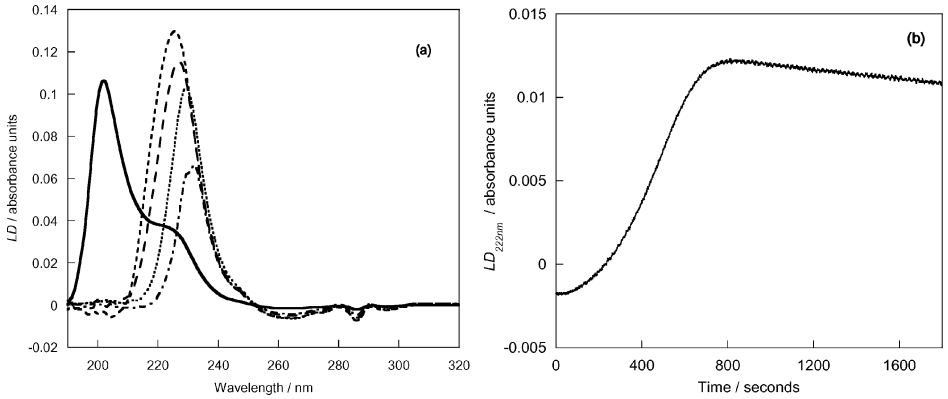


Fig. 7 (a) Far UV LD spectra showing the apparent shift to shorter wavelength of the maximum signal as the concentration of F-actin is reduced. F-actin concentrations 93; 74; 62; 53 and 12 μM (the true spectrum, solid line). (b) Polymerisation of G-actin into F-actin (~0.4 mg mL^{-1}) with ATP present (a) kinetics analysis monitoring LD_{222nm}.

orientations of aromatic residues in actin and the actin–myosin complex.

Tropomyosin is another actin-binding protein, but in this case it is itself a linear molecule and it binds to actin with a low stoichiometry of A_7Tm. Tropomyosin (Tm) is a coiled-coil protein that aligns in Couette flow and has its own LD spectrum, as first reported in ref. 47 where data were collected with a large-volume LD cell down to ~210 nm. The LD spectrum of Tm (Fig. 9) has a broad negative LD signal between 218–240 nm (n → π* transition) and a positive LD maxima at 206 nm.[46] Below 206 nm the LD signal tends negative. At higher concentrations a positive LD signal in the near UV at 280 nm due to the tyrosines (Tm contains no tryptophans) is apparent. Tm is a highly helical protein (>90% α-helices[48]), and the n → π* transition of the α-helices is perpendicular to the orientation axis, so its negative LD is consistent with the α-helices being more parallel than perpendicular to the orientation axis, as one would expect for a coiled-coil. The presence of Tm bound to actin has little effect on the LD spectrum of actin, largely because of the weak Tm signal. Tm forms a 1 : 1 complex with troponin (Tn, a globular protein which therefore has no intrinsic flow LD signal) which then binds to actin in a 7 : 1 : 1 ratio. Fig. 10 shows LD spectra of phalloidin-stabilised actin (psA, to make extended data collection feasible), Tm and Tn individually, psA-Tm and psA-Tn mixtures and in a 1 : 1 : 1 mixture to ensure saturation of actin with Tm and Tn. Tn increases the psA signal slightly, Tm shifts its maximum wavelength and also leads to a slightly increased signal, whereas the 1 : 1 : 1 mixture spectrum has both a shift of wavelength and a 30% larger LD signal. These data indicate that Tn's main effect is on the orientation factor of the fibre. Upon addition of $CaCl_2$, the LD spectrum (of psA–Tm–Tn–Ca^{2+}) overlays that of psA–Tm consistent with calcium causing the Tn–actin interaction to weaken (data not shown). Near UV LD studies by Yanagida *et al.* in 1974 on the effect of Ca^{2+} on the F-actin–Tm–Tn complex showed about a 20% decrease in intensity of the two negative peaks at 260 and 295 nm, and concluded that the presence of calcium made the fibre more flexible.[24,45] Taniguchi reported an even larger decrease in LD intensity upon the addition of calcium ions.

Near UV LD of actin (21 μM) and its complex with Tm (3 μM), Tm–Tn (3 μM) and Tm–Tn–Ca^{2+} (Tm–Tn 3 μM and Ca^{2+} 3 mM) are shown in Fig. 11.[46] It can be seen that there is an increase in signal intensity of the 285 nm transition (tyrosine and tryptophan) upon binding of Tm, and even

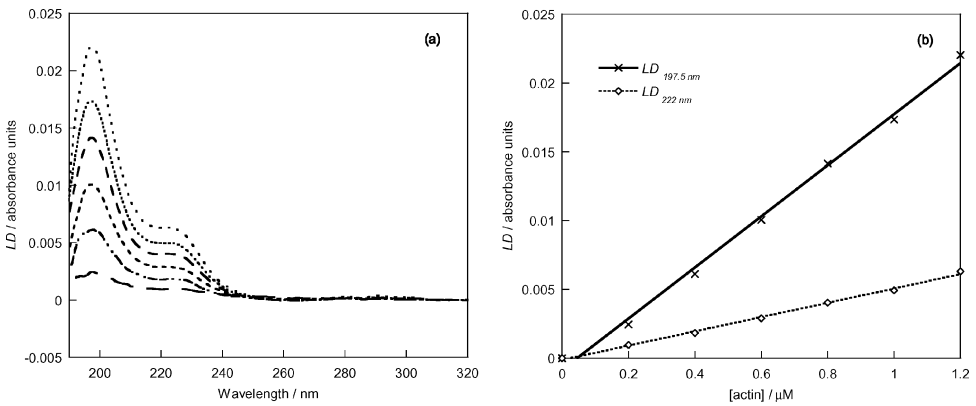


Fig. 8 The concentration dependence of the LD of solutions of F-actin at low concentrations, (a) wavelength LD spectra with F-actin concentrations of 1.2 (···); 1.0 (·····); 0.8 (– – –); 0.6 (- - - -); 0.4 (· · – – –) and 0.2 μM (— · —), and (b) linear plots showing that the Beer Lambert law is obeyed at the concentrations shown.

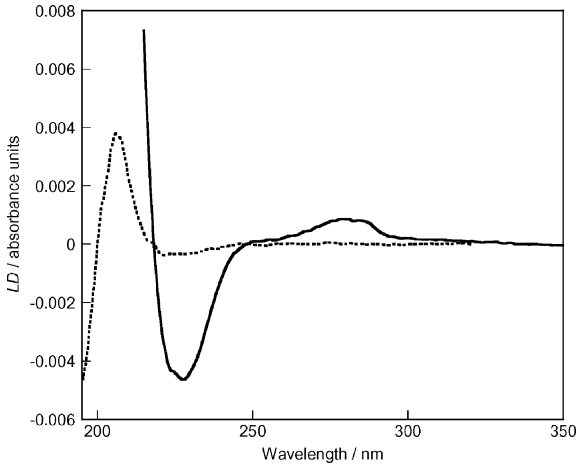


Fig. 9 LD spectra of tropomyosin at 50 (——) and 10 μM (- - - -) in KCl (20 mM), $MgCl_2$ (1 mM) and MOPS (4 mM) pH 7.

more so (by ~30%) upon binding of the Tm–Tn complex. The ADP region (265 nm) shows a slight increase upon addition of Tm to actin, with no noticeable change recorded for the addition of Tm–Tn to actin. There is no obvious difference in the tryptophan signal at 295 nm except for the binding of the Tm–Tn complex to actin. The effect of the addition of calcium to the system reduced all three signals by ~30%. This is consistent with literature data, though no change in the 285 nm transition had previously been reported. This may be due to the methods employed for the correction of scattering when these analyses were originally undertaken. As concentrations of greater than 100 μM were used in 1970s, light scattering artefacts were very significant in the spectra collected.

When using linear dichroism to detect the binding between a fibrous protein and globular protein (such as actin and myosin) it can be difficult to determine whether changes in the LD are due to the globular protein once it is oriented and/or whether there is a change in the orientation of the fibre as a whole or even its subunits. To aid in the interpretation of such

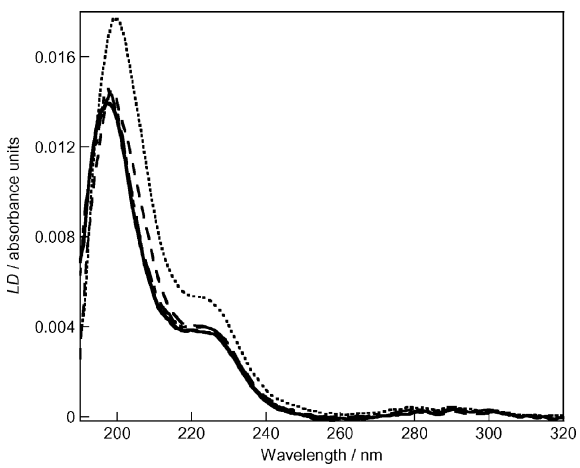


Fig. 10 (a) LD of phalloidin stabilised actin (psA) (——), psA–Tn (– – –); psA–Tm (- - - -); psA–Tm–Tn (. . ...). All protein concentrations are 1 μM. All solutions contain KCl (15 mM); $MgCl_2$ (0.75 mM) and MOPS (3 mM) pH 7.

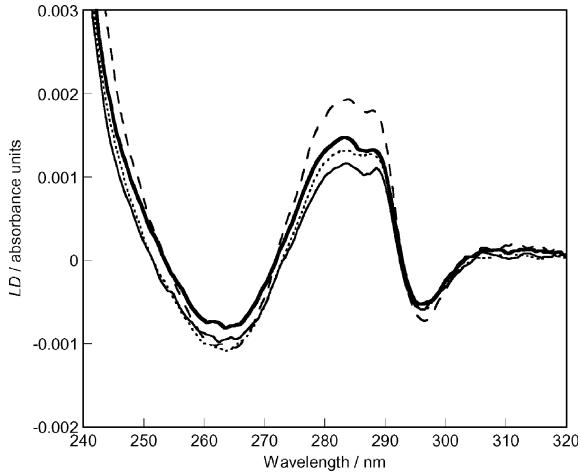


Fig. 11 Near UV LD spectra of actin (A) (21 μM) (——); A (21 μM)–tropomyosin (Tm) (3 μM) (. . . . .); A (21 μM)–Tm (3 μM)–Troponin (Tn) (3 μM) (- - - -) and A (21 μM)–Tm (3 μM)–Tn (3 μM)–Ca^{2+} (3 mM) (——). All solutions contained KCl (15 mM); $MgCl_2$ (0.75 mM) and MOPS (3 mM) pH 7.

spectra a probe molecule attached to the fibre that allows the orientation factor to be evaluated is required. The ideal probe has yet to be found, but Fig. 12 shows that probes can be readily detected. Rhodamine labelled phalloidin is attractive as the phalloidin stabilises the actin filaments without affecting the structure (rhodamine-labelled phalloidin stabilised actin and actin have very similar shaped LD spectra in the UV despite contributions from the ligand) and rhodamine has a large extinction coefficient (106 000 M^{-1} cm^{-1} at 542.75 nm[49]) and the signals are linear with concentration.

Far UV LD of other fibrous proteins

As noted above, actin was the first protein fibre to which Couette flow *LD* was applied in the near UV wavelength region and it is now possible to collect far UV data on such a system. A similar story can be told for tubulin (where α,β-tubulin heterodimers nucleate into oligomers, which then form protofilaments, with the subunits aligned head to tail with dimeric units repeating every 8 nm). Early experiments probed the near UV[38] or used the turbidity LD signal outside the absorbance region to follow polymerisation.[39] Backbone data for tubulin have more recently been collected and used to monitor chromophore reorientation during the polymerisation process.[5] One of the challenges of tubulin is the need to hold the sample at 37 °C during polymerisation. Analysis of tubulin kinetics and binding of various ligands including Taxol™, colchicine, vincristine and DAPI have been probed in a microvolume Couette cell.[5]

Essentially any fibrous protein which stays in solution can be flow oriented; at least preliminary near and far UV LD data have been collected for prions,[50] Alzheimer's fibres, collagen, and α_1-antitrypsin.[11] LD has also been used to probe the processes involved in the polymerisation and bundling of the bacterial homologue of tubulin, FtsZ.[51] The *E. coli* FtsZ has no tryptophans so the near UV region of the spectrum can be

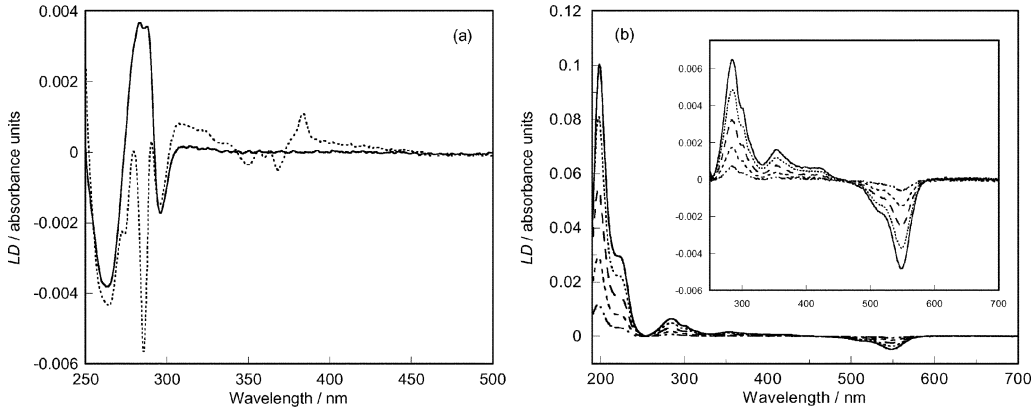


Fig. 12 (a) LD of actin (50 μM) (——) and pyrene labelled actin (50 μM) (· · · ·). (b) LD of rhodamine phalloidin labelled actin 10 (——), 8 (· · · ·), 6 (– – –), 4 (- - -) and 2 μM (· · – –) showing linear increase of LD signal with concentration.

interpreted in terms of contributions from guanine and tyrosines. This has been used to follow some of the structural changes which occur when the protofilaments bundle to form fibres. LD can also be used to probe the binding of another protein to a fibre. For example, the *E. coli* ortholog of the accessory protein Zap A, YgfE (which has been shown to be a bonafide division protein[52]) has been shown to promote FtsZ bundling at physiological concentrations.[52] During the bundling process LD shows that the guanine in GTP (which is sandwiched between successive FtsZ units) changes orientation. This has led to the proposal that macrochelation by a divalent metal cation to the terminal phosphate of GTP and the N7 nitrogen of guanine causes the conformational change of the guanine. However, whether this initiates or results from FtsZ polymer bundling is at present unknown.

DNA and DNA–ligand systems

An article with the chosen title would not be complete without at least some mention of DNA and DNA–ligand systems, due to the extensive literature available in this area. However, this also presents a problem as it is very difficult to give a representative view of the whole field. B-DNA is a long molecule whose base π–π* transitions all lie perpendicular to the helix axis in an idealised structure—thus giving rise to a negative signal under the DNA absorbance bands with a maximum magnitude at the same wavelength (almost) as the normal absorbance maximum. The consensus is that rather than 90°, a value of 86° should be used (though the 'true' value is probably less than this and depends on the flexibility (so environment and sequence) of a given piece of DNA). Chou and Johnson[53] extensively analysed the LD of DNA down to 175 nm in terms of the base transitions and found the average inclination of the bases is of the order of 20° from the helix axis.

One of the main applications of LD uses its ability to probe interactions between an oriented molecule and a species that can not be aligned by itself but aligns upon binding to the long molecule. Information regarding the corresponding binding geometries can be ascertained if transition moment directions of the components are known, or conversely. Molecules (usually cationic species) can bind to DNA by intercalation, groove binding or externally binding. Examples include dyes, metal complexes and proteins. There are a number of superb reviews on this area, as noted above.[6,7] The work of Schellman and collaborators and Nordén and collaborators has been particularly valuable. Table 1 gives examples of DNA–ligand complexes studied by LD.

When ligands are added to a solution of DNA and they bind, one expects to see signals due to the ligand transitions as well as the DNA. Ideally the ligand has no absorbance in the DNA region, so one can determine the DNA orientation parameter in the presence of the ligand. Intercalators are planar aromatic molecules and bind parallel to the DNA bases, so to a first approximation are expected to have LD^r values the same as the DNA bases. In practise, they locally stiffen the DNA so they and the bases near them are better oriented than other bases and the LD^r is greater in the ligand region.[8] Thus, if all ligand transitions have negative LD signals with LD^r values greater (due to increase in S) or equal in magnitude to those of the DNA bases, then one can be fairly sure that the ligand is intercalating. Alternatively, if long-axis polarised ligand transitions have positive LD signals, then a groove-binding mode is indicated. If there is a preference for AT-rich regions of DNA (where the groove is less sterically hindered) this supports this binding mode assignment.[54] If S is known, one can be more quantitative above this conclusion. If

Table 1 Examples of DNA–ligand complexes studied by LD. This table together with ref. 6 and 7 gives the beginning of a literature trail to find most systems which have been studied

Type of compound bound to DNA	Ligand bound to DNA	Ref.
Dyes	Ethidium Bromide	58
	DAPI	59
	Hoechst	60
	9-Hydroxyellipticine	61
Metal complexes	General	62
	Ruthenium metal complexes	63
		64
	Di-iron metal complexes	56
	Cobalt ammines	65
Proteins	Chromatin	66
	Rec-A	67

the only available ligand transition is in the DNA region one either has to hope S has not changed or hope that there is a wavelength where DNA can be probed independently of the ligand, as was the case for anthracene deriviatised with cationic spermine.[55] In this case a slight bending or stiffening of the DNA was observed. In the case of a tetracationic major groove binding di-iron helicate, the DNA was significantly bent locally to the ligand binding site, so the ligand orientation on the DNA cannot be determined. However, at low loading the bending per ligand was calculated.[56] An additional challenge in quantifying ligand orientation is that only bound ligands contribute to the LD signal but unbound ones contribute to the absorbance, thus making LD^{r} determinations challenging if the binding constant and the different extinction coefficients of free and bound ligand are not known.[57]

The extent of orientation of a DNA molecule depends on its flexibility and also its length. Simonson and Kubista[68] empirically related LD intensity to DNA length on the basis of the LD^{r} of some long DNAs of well-defined length:

$$LD^{r} = \frac{LD}{A} = -\frac{k_1 G}{k_2 + G} \qquad (7)$$

where A is the absorbance, G is the velocity gradient of the flow cell, $k_1 = 0.42$ for our experiment (from Fig. 8 of ref. 68), and the inverse of $1/k_2$ is proportional to the DNA length.

This equation has since been used to quantify the bending induced in DNA upon addition of a ligand.[56] The minimum length of DNA required to give an LD signal depends on the DNA flexibility and the desired signal : noise ratio. For easy data collection, 1000 or more base pairs are desirable. Less than 250 base pairs will be hard work—not surprisingly given that the persistence length (the length of DNA for it to bend though 1 rad) of DNA is typically $\sim$150 base pairs. Less quantitatively, changes in LD can be used to monitor single nucleotide polymorphisms since DNA mis-matches have an effect on DNA shape (hence scattering) and DNA orientation.[69]

Using molecular length to follow kinetics

LD is the ideal method to probe the kinetics of formation of protein fibres from assemblies of monomeric units as discussed above. It is also the ideal technique to follow such reactions as the effect of a restriction enzyme (DNA-cutting enzyme) as a function of time.[70] As illustrated in Fig. 13 if one begins with comparatively compact supercoiled circular DNA and adds a restriction enzyme the LD signal increases as the average length and hence orientation of the DNA increases. If the enzyme has two cut sites, then a subsequent decrease in LD is noticed as the average length decreases again. Perhaps the most intriguing thing about this experiment is that a supercoiled plasmid DNA (in this case $\sim$7000 base pairs in length) can be readily oriented—the starting point is not a zero LD signal. It is conceivable that, coupled with appropriate modelling, LD may be able to give information about the structures adopted by such a molecule.

LD is also ideal for following the progress of polymerase chain reactions (PCRs) or even simply monitoring the end point of a reaction.[71] In the latter case it can be used to determine viral load and by careful choice of primers this can be selective. The advantage of LD over other real time PCR methods is that primer-dimers, which plague most quantitative PCR methods, are invisible. Also, no additional labelling is required.

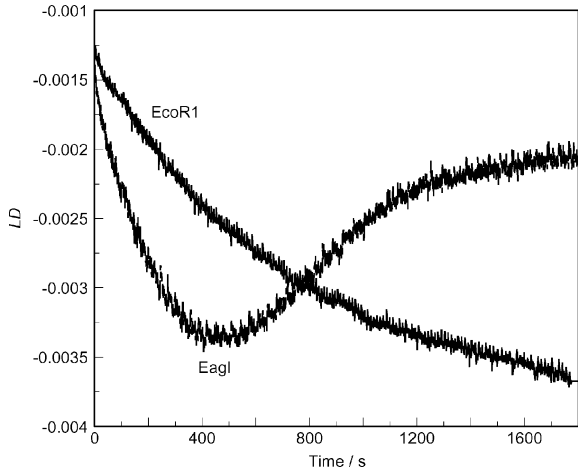


Fig. 13 Kinetics of restriction digests of a circular, super-coiled plasmid with *Eco*RI (which has a single cut site in the DNA sequence and thus single step kinetics) and *Eag*I (which has two cut sites in the DNA sequence).

Carbon nanotubes

Carbon nanotubes provide a final example for this article and 'proof' that one can use flow LD for any long molecule which can be put into solution and is spectroscopically active. Single walled carbon nanotubes (SWNTs) are soluble (more or less) in aqueous SDS. Despite the size of the SWNTs the scattering does not dominate the LD signal and the nanotube itself has a broad band in the far UV region of the spectrum.[72,73] Any small molecule which binds to DNA also acquires an LD signal, though in the case of aromatic molecules it is clear that there is significant coupling between the graphene sheets of the nanotubes and the π-systems of the ligand as well as inter-ligand interaction (as shown by the concentration dependence of the shape of the spectrum in Fig. 14).[70] Despite the negative charge of DNA we have also oriented DNA and the neutral PNA on SWNTs. The signs and magnitudes of the DNA signals should enable an accurate average orientation of the DNA on the SWNT to be determined if we know the orientation parameter. So far only qualitative analyses have been possible which indicate that DNA wraps around the nanotube in the opposite manner from PNA.[73]

Conclusion

Other recent applications of LD include the analysis of proteins and peptides in lipid bilayers, for example cytochrome *c* inserted into soybean liposome,[15] and the orientation of peptide fibres that are designed for molecular self assembly.[74] LD has also successfully been used to provide experimental evidence for models postulated by computer

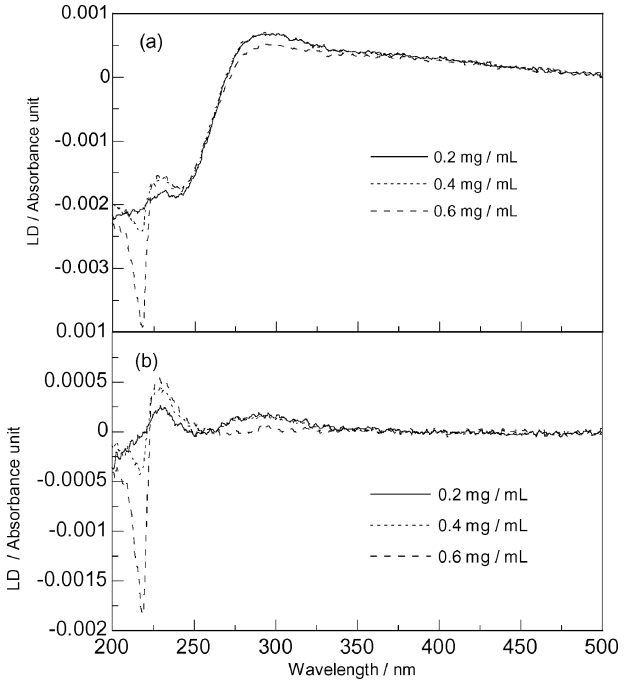


Fig. 14 (a) LD spectra of different concentrations of naphthalene in SWNTs (0.1 mg mL^{-1}) in SDS (9 mM) solution. SWNTs–naphthalene complexes were prepared by adding different amounts (0.2, 0.4 and 0.6 mg) of methanol dissolved naphthalene to 1 mL of SWNTs (0.1 mg mL^{-1}) in SDS (9 mM) solution and making final naphthalene concentrations of 0.2, 0.4 and 0.6 mg mL^{-1}. (b) Difference spectra obtained by subtracting the LD spectrum of SWNT (0.1 mg mL^{-1}) from the LD spectra of 0.2, 0.4 and 0.6 mg mL^{-1} naphthalene–SWNTs complexes.

simulations.[64] An ongoing project is the use of computer modelling programs for further structural characterisation of LD spectra of protein fibres.[15] So in conclusion it really does seem as if flow linear dichroism can be used to study any long molecule whose absorption spectrum is accessible.

Appendix: Materials and methods for actin studies

Unlabelled and pyrene-labelled actin were prepared using the methods detailed in ref. 75 and 76. Phalloidin-stabilised F-actin was prepared by incubating a solution of 10 μM actin (unlabelled or labelled) with 10 μM phalloidin for at least one hour at 4 °C.[77] Rhodamine phalloidin (R-415) was purchased from Molecular Probes. Using the information provided on the data sheet, a concentration of ~200 μM was prepared by dissolving the contents of the vial in 49.5 μL methanol (spectroscopic grade). Rhodamine-labelled phalloidin-stabilised actin was prepared by mixing rhodamine-labelled phalloidin and unlabelled actin in a 1 : 1 ratio at 10 μM for at least one hour at 4 °C. Tropomyosin (Tm) and troponin (Tn) were purchased from Sigma as lyophilized powder and solutions were prepared on the day of analysis. A relative molecular mass of ~65 kDa was assumed for Tm (calculated using Swiss-Prot accession numbers P04268 and P19352 for the α and β chains of tropomyosin from chicken[78]) with a purity of 99% (as stated on the container). A relative molecular mass of ~73.0 kDa was assumed for Tn (using Swiss-Prot accession numbers P02644; P12620 and P09860 for each subunit of Tn[78]), with a purity of 91% (as stated on container).

References

1 R. Marrington, T. R. Dafforn, D. J. Halsall and A. Rodger, *Biophys. J.*, 2004, **87**, 2002–2012.
2 A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Nordén, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez and M. J. Hannon, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4051–4057.
3 A. Rodger, J. Rajendra, R. Marrington, R. Mortimer, T. Andrews, J. B. Hirst, A. T. B. Gilbert, D. Halsall, T. Dafforn, M. Ardhammar, B. Nordén, C. A. Woolhead, C. Robinson, T. Pinheiro, K. J. M. Seymour, N. Perez and M. J. Hannon, in *Biophysical Chemistry: Membranes and Proteins*, ed. R. H. Templer and R. Leatherbarrow, The Royal Society of Chemistry, Cambridge, 2002, pp. 3–19.
4 R. Marrington, T. R. Dafforn, D. J. Halsall, M. Hicks and A. Rodger, *Analyst*, 2005, **130**, 1608–1616.
5 R. Marrington, M. Seymour and A. Rodger, *Chirality*, 2006.
6 B. Nordén, M. Kubista and T. Kuruscev, *Q. Rev. Biophys.*, 1992, **25**, 51–170.
7 B. Nordén, *Appl. Spectrosc. Rev.*, 1978, **14**, 157–248.
8 A. Rodger and B. Nordén, *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, 1997.
9 B. Albinsson and B. Nordén, *J. Phys. Chem.*, 1992, **96**, 6204–6212.
10 A. Holmen, A. Broo, B. Albinsson and B. Norden, *J. Am. Chem. Soc.*, 1997, **119**, 12240–12250.
11 T. R. Dafforn, J. Rajendra, D. J. Halsall, L. C. Serpell and A. Rodger, *Biophys. J.*, 2004, **86**, 404–410.
12 A. Rodger, in *Methods in Enzymology*, ed. J. F. Riordan and B. L. Vallee, Academic Press, San Diego, 1993, vol. 226, pp. 232–258.
13 T. R. Dafforn, J. Rajendra, D. J. Halsall, L. C. Serpell and A. Rodger, *Biophys. J.*, 2004, **86**, 404–410.
14 L. B. A. Johansson and A. Davidsson, *J. Chem. Soc., Faraday Trans. 1*, 1985, **81**, 1375–1388.
15 A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Norden, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. T. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez and M. J. Hannon, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4051–4057.
16 M. Couette, *Ann. Chim. Phys*, 1890, **6**, 433–510.
17 A. Mallock, *Proc. R. Soc. London*, 1888, **45**, 126.
18 A. Mallock, *Philos. Trans. R. Soc. London, Ser. A*, 1896, **187**, 41.
19 M. Kasai and F. Oosawa, *Methods Enzymol.*, 1972, **25**, 289–323.
20 A. Wada and S. Kozawa, *J. Polym. Sci., Part A*, 1964, **2**, 853–864.
21 S. Higashi, M. Kasai, F. Oosawa and A. Wada, *J. Mol. Biol.*, 1963, **7**, 421–430.
22 J. Hofricheter and W. Eaton, *Annu. Rev. Biophys. Bioeng.*, 1976, **5**, 511–560.
23 B. Nordén, *Appl. Spectrosc. Rev.*, 1978, **14**, 157–248.
24 F. Oosawa, Y. Maeda, S. Fujime, S. Ishiwata, T. Yanagida and M. Taniguchi, *J. Mechanochem. Cell Motil.*, 1977, **4**, 63–78.
25 B. Nordén, M. Kubista and T. Kurucsev, *Q. Rev. Biophys.*, 1992, **25**, 51–170.
26 R. J. Donnelly, *Phys. Today*, 1991, **November**, 32–39.
27 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1923, **223**, 289–343.
28 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1936, **157**, 546–564.
29 G. I. Taylor, *Proc. R. Soc. London, Ser. A*, 1936, **157**, 565–578.
30 A. Wada, *Biopolymers*, 1964, **2**, 361–380.
31 C. Lee and N. Davidson, *Biopolymers*, 1968, **6**, 531–550.
32 P. Oriel and J. Schellman, *Biopolymers*, 1966, **4**, 469–494.
33 J. O. Wilkes, *Fluid Mechanics for Chemical Engineers*, Prentice Hall PTR, Upper Saddle River, 1999.
34 C. D. Anderick, S. S. Liu and H. L. Swinney, *J. Fluid Mech.*, 1986, **164**, 155–183.
35 O. Reynolds, *Philos. Trans. R. Soc. London*, 1883, **174**, 935–982.
36 O. Reynolds, *Proc. R. Soc. London*, 1883, **35**, 84–99.
37 M. C. Potter and D. C. Wiggert, *Mechanics of Fluids*, Brooks/Cole Thomson Learning, Pacific Grove, 3rd edn, 2002.

38 M. Taniguchi and R. Kuriyama, *Biochim. Biophys. Acta*, 1978, **533**, 538–541.
39 J. Nordh, J. Deinum and B. Norden, *Eur. Biophys. J.*, 1986, **14**, 113–122.
40 E. D. Korn, M. F. Carlier and D. Pantaloni, *Science*, 1987, **238**.
41 S. L. Brenner and E. D. Korn, *J. Biol. Chem.*, 1984, **259**, 1441–1446.
42 K. C. Holmes, D. Popp, W. Gebhard and W. Kabsch, *Nature*, 1990, **347**, 44–49.
43 M. Miki and K. Mihashi, *Biophys. Chem.*, 1977, **6**, 101–106.
44 M. Taniguchi, *Electro–Opt. Dielectr. Macromol. Colloids*, 1979, 203–210.
45 T. Yanagida, M. Taniguchi and F. Oosawa, *J. Mol. Biol.*, 1974, **90**, 509–522.
46 R. Marrington, PhD Thesis, University of Warwick (Coventry), 2004.
47 M. J. Pandya, G. M. Spooner, M. Sunde, J. R. Thorpe, A. Rodger and D. N. Woolfson, *Biochemistry*, 2000, **39**, 8728–8734.
48 F. G. Whitby and G. N. Phillips, Jr, *Proteins: Struct., Funct., Genet.*, 2000, **38**, 49–59.
49 Molecular Probes, UK, 1993.
50 L. C. Serpel, 2003, unpublished data.
51 R. Marrington, E. Small, A. Rodger, T. R. Dafforn and S. Addinall, *J. Biol. Chem.*, 2004, **47**, 48821–48829.
52 F. J. Gueiros-Filho and R. Losick, *Genes Dev.*, 2002, **16**, 2544–2556.
53 P.-J. Chou and W. C. J. Johnson, *J. Am. Chem. Soc.*, 1993, **115**, 1205–1214.
54 B. Nordèn, S. Eriksson, S. K. Kim, M. Kubista, R. Lyng and B. Akerman, in *The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, ed. B. Pullman and J. Jornter, Kluwer Academic Publishers, Dordrecht, 1990, vol. 23, pp. 23–41.
55 A. Rodger, I. S. Blagbrough, G. Adlam and M. L. Carpenter, *Biopolymers*, 1994, **34**, 1583–1593.
56 M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chem., Int. Ed.*, 2001, **40**, 879–884.
57 D. Z. Coggan, I. S. Haworth, P. J. Bates, A. Robinson and A. Rodger, *Inorg. Chem.*, 1999, **38**, 4486–4497.
58 C. Houssier, B. Hardy and E. Fredercq, *Biopolymers*, 1974, **13**, 1141–1160.
59 M. Kubista, B. Akerman and B. Norden, *Biochemistry*, 1987, **26**, 4545–4553.
60 J. H. Moon, S. K. Kim, U. Sehlstedt, A. Rodger and B. Norden, *Biopolymers*, 1996, **38**, 593–606.
61 M. A. Ismail, K. Sanders, G. C. Fennell, H. C. Latham, P. Wormell and A. Rodger, *Biopolymers*, 1998, **46**, 127–143.
62 B. Nordén and F. Tjerneld, *FEBS Lett.*, 1976, **67**, 368–370.
63 K. K. Patel, E. A. Plummer, M. Darwish, A. Rodger and M. J. Hannon, *J. Inorg. Biochem.*, 2002, **91**, 220–229.
64 D. Z. M. Coggan, I. S. Haworth, P. J. Bates, A. Robinson and A. Rodger, *Inorg. Chem.*, 1999, **38**, 4486–4497.
65 A. Rodger, A. Parkinson and S. Best, *Eur. J. Inorg. Chem.*, 2001, **9**, 2311–2316.
66 F. Tjerneld, B. Nordén and H. Wallin, *Biopolymers*, 1982, **21**, 343–358.
67 B. Nordén, K. Mortensen, C. Elvingson, M. Kubista, B. Sjöberg, M. Ryberg and M. Takahashi, *J. Mol. Biol.*, 1991, **226**, 1175–1191.
68 T. Simonson and M. Kubista, *Biopolymers*, 1993, **33**, 1225–1235.
69 T. R. Dafforn, D. J. Halsall and A. Rodger, *Chem. Commun.*, 2001, 2410–2411.
70 M. R. Hicks, A. Rodger, C. M. Thomas, S. Blatt and T. R. Dafforn, *Biochemistry*, 2006, submitted for publication.
71 D. J. Halsall, T. R. Dafforn, R. Marrington, E. Halligan and A. Rodger, *IVD Technol.*, 2004, **6**, 51–60.
72 J. Rajendra, M. Baxendale, L. G. Dit Rap and A. Rodger, *J. Am. Chem. Soc.*, 2004, **126**, 11182–11188.
73 J. Rajendra and A. Rodger, *Chem.–Eur. J.*, 2005, **11**, 4841–4848.
74 M. J. Pandya, G. M. Spooner, M. Sunde, J. R. Thorpe, A. Rodger and D. N. Woolfson, *Biochemistry*, 2000, **39**, 8728–8734.
75 S. S. Lehrer and G. Kerwar, *Biochemistry*, 1972, **11**, 1211–1217.
76 A. H. Criddle, M. A. Geeves and T. E. Jeffries, *Biochem. J.*, 1985, **232**, 343–349.
77 R. Maytum, S. S. Lehrer and M. A. Geeves, *Biochemistry*, 1999, **38**, 1102–1110.
78 E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch, *Nucleic Acids Res.*, 2003, **31**, 3784–3788.

Improved curve fitting procedures to determine equilibrium binding constants

Frank H. Stootman,*[a] Dianne M. Fisher,[b] Alison Rodger[c] and Janice R. Aldrich-Wright*[a]

Received 31st March 2006, Accepted 18th July 2006
First published as an Advance Article on the web 7th August 2006
DOI: 10.1039/b604686j

For ligand–biomacromolecule titration experiments it has been traditional practice to extract parameters such as the equilibrium binding constant K and the number of bases per ligand binding site n with relatively labour intensive methods, usually based on single wavelength data, such as the difference method by Rodger and Nordén coupled together with a Scatchard plot. Presented in this paper are both the theory and a least squares fitting method to derive parameters such as K and n more directly from all spectral non-linear experimental data. Both the case of non competitive binding of a metal complex ligand to DNA and the case of displacement by a metal complex ligand of an ethidium marker attached to the DNA are considered. This work may be applied directly to reduce experimental data produced by a spectropolarimeter (for circular or linear dichroism) or a spectrophotometer (for fluorescence or UV-Vis spectroscopy).

1 Introduction

Equilibrium binding experiments are widely used to investigate interactions between small molecules and larger biological macromolecules such as DNA.[1,2,5,9,10,12,17,19] Interactions between ligands and DNA may involve binding modes such as intercalation, H-bonding, van der Waals groove binding or *pi* stacking along the outside of the double helix. The determination of the equilibrium rate constant K at which such interactions occur make it possible to evaluate and compare the binding strength (or binding affinity) of a number of related drug compounds for a particular macromolecule under specific experimental conditions.

Over the years, equilibrium binding constants K have been calculated using a variety of methods developed by Scatchard,[23] McGhee and von Hippel,[11] Nordén and Tjerneld,[13] Wolfe *et al.*,[24] Rodger,[20] Rodger and Nordén,[22] and Kumar and Asuncion[7] among many others. Many of the traditional techniques use single wavelength data and require numerous steps to determine K and n (the number of bases per binding site). They are often time consuming, involving operator judgement and intermediate steps.

In this paper we present an improved method for the determination of the equilibrium binding constant which is obtained directly from all of the experimental data using a non-linear least squares fit of a simple equilibrium model also used by Rodger[20] and Rodger and Nordén[22] and their co-workers in their, so called, Intrinsic Method for the case of non-competitive metal complex ligand binding to DNA. We also present a significant advance to include the case of competitive binding in which a fluorescent marker molecule such as ethidium which is bound to the DNA is displaced by the metal complex.

[a]*University of Western Sydney, PO Box 1797, Penrith South DC, NSW 1797, Australia. E-mail: f.stootman@uws.edu.au*
[b]*The University of Sydney, Centre for Heavy Metals Research, School of Chemistry, The University of Sydney, NSW 2006, Australia*
[c]*The University of Warwick, Coventry, UK CV4 7AL.*
E-mail: j.aldrich-wright@uws.edu.au; Fax: +612 4620 3025

2 Non-competitive binding theory for DNA–ligand systems

We shall use the following symbol convention for titration of metal complex into a fixed concentration of DNA in buffer or *vice versa*: L_F to represent free metal complex; S_F to represent free binding sites on the DNA; and L_B to represent bound metal complex sites. We assume that L_B, bound to the DNA, can be measured using macroscopic techniques such as a spectropolarimeter (for CD or LD) or a spectrophotometer (for fluorescence or UV-Vis). The equilibrium rate equation for the titration is as follows:

$$L_F + S_F \underset{}{\overset{K}{\rightleftharpoons}} L_B. \quad (1)$$

The equilibrium rate constant is then simply,

$$K = \frac{L_B}{L_F S_F} \quad (2)$$

Let L_T be the total number of metal complexes, then, $L_T = L_F + L_B$. Similarly let S_T be the total number of ligand binding sites on the DNA, then, $S_T = S_F + S_B$. Substituting into eqn (2) and making the reasonable assumption that $L_B = S_B$, we obtain,

$$K = \frac{L_B}{(L_T - L_B)(S_T - L_B)} \quad (3)$$

Re-arranging eqn (3) and making L_B the subject of the formula, we obtain the quadratic,

$$L_B^2 - L_B\left(\frac{1}{K} + L_T + S_T\right) + S_T L_T = 0. \quad (4)$$

The standard general solution of eqn (4) is,

$$L_B = \frac{1}{2}\left\{\frac{1}{K} + L_T + S_T - \sqrt{\left(\frac{1}{K} + L_T + S_T\right)^2 - 4S_T L_T}\right\} \quad (5)$$

This equation is similar to that reported and used by Pilch *et al.*[15]

It is important to note that there are actually *two* solutions for L_B in eqn (4). Thus choice of the negative sign before the square root in eqn (5) is governed by the need to have a physically realistic solution. The solution is also symmetric in L_T and S_T. It can be readily shown that if L_T is variable and S_T a constant, then $L_B \rightarrow S_T$, and if S_T is a variable and L_T a constant, then $L_B \rightarrow L_T$. Both asymptotes are expected depending on experimental design.

The bound metal complex L_B has an experimental measurement ε which is related to L_B *via* an instrumental response sensitivity R_b. The experimental curve is therefore related to the above theory by:

$$\varepsilon = R_b L_B \quad (6)$$

3 Competitive binding theory—displacement experiments

One method of determining the binding of an invisible (or not sufficiently visible) ligand is to measure the change in signal of another molecule which the ligand of interest displaces. For example, one can add ethidium bromide to DNA and then displace it from the DNA with another ligand such as a metal complex. If the metal complex is not fluorescent then the displacement of the ethidium from the DNA causes a decrease in the measured fluorescence. This information may be used to obtain the equilibrium binding constant of the non-fluorescent metal complex as outlined below.

The reaction equations for this scenario are given in eqn (7) and (8).

$$E_F + S_F \underset{}{\overset{K_E}{\rightleftharpoons}} E_B, \quad (7)$$

$$L_F + S'_F \underset{}{\overset{K_L}{\rightleftharpoons}} L_B. \quad (8)$$

In order to deal with the different number of base pairs associated with the ethidium binding (n_E) and metal complex binding (n_L) we have designated the free binding sites as $S_F = B_F/n_E$ and $S'_F = B_F/n_L$, respectively, where B_F is the number of free base pairs on the DNA. The equilibrium binding constants are,

$$K_E = \frac{n_E E_B}{E_F B_F} \quad (9)$$

$$K_L = \frac{n_L L_B}{L_F B_F} \quad (10)$$

To demonstrate the relationship of eqn (7–10) to the literature, let r be the simple ratio: moles bound ethidium/total moles of DNA binding sites with respect to ethidium. It follows,

$$r = \frac{E_B}{S_T} = \frac{E_B}{S_F + (n_L/n_E)L_B + E_B}. \quad (11)$$

Using eqn (9) and (10), and introducing K, defined as follows,

$$K = \frac{K_E}{1 + K_L L_F}, \quad (12)$$

leads to the classic formulation given in the literature,[4,6,8]

$$\frac{r}{E_F} = K(1-r), \quad (13)$$

Eqn (13) is converted to a quadratic form starting from eqn (11). We make E_B the subject of the formula because this is the experimental marker for the progress of its displacement by the metal complex. Then,

$$E_B^2 - E_B\left(\frac{1}{K} + S_T + E_T\right) + S_T E_T = 0. \quad (14)$$

This has a similar form to eqn (4). In the competitive binding case, however, K is dependent on L_F, which is a variable and an unknown quantity as the metal complex titration progresses in the displacement reaction. Nevertheless, it is always possible to write: $aL_B + E_B = S_B$, where $a = n_L/n_E$. Hence,

$$L_F = L_T - L_B = L_T - \frac{S_B - E_B}{a}, \quad (15)$$

It follows,

$$\frac{1}{K} = \frac{1 + K_L(L_T - S_B/a)}{K_E} + \frac{K_L E_B}{K_E a}, \quad (16)$$

which when substituted into eqn (14) leads to a slightly more complex quadratic expression,

$$E_B^2\left(1 - \frac{K_L}{aK_E}\right) - E_B\left(\frac{K_L}{K_E}L_T - \frac{K_L}{aK_E}S_B + \frac{1}{K_E} + S_T + E_T\right) + S_T E_T = 0. \quad (17)$$

Initially displacement experiments need to be done in two steps. In the first step, there is no metal complex ($L_F = 0$) and the ethidium is titrated into the DNA (or *vice versa*). In this case $K = K_E$ and eqn (17) reduces to the form of eqn (4) and its solution can be treated in the manner described in Section 4.1. As a result of the first step the following constants can be set before metal complex titration is commenced:

$$\alpha = \frac{1}{K_E} + S_T + E_T \quad (18)$$

$$\beta = S_T E_T \quad (19)$$

In the second step, metal complex is titrated to displace the ethidium. The theoretical equation is simplified to,

$$E_B^2\left(1 - \frac{K_L}{aK_E}\right) - E_B\left(\frac{K_L}{K_E}L_T - \frac{K_L}{aK_E}S_B + \alpha\right) + \beta = 0. \quad (20)$$

In order to obtain a physically correct asymptotic solution in which $E_B \rightarrow 0$ as $L_T \rightarrow \infty$ we use the positive root in the quadratic solution to the above equation for the case $K_L/aK_E > 1$. The negative root is for the case $0 < K_L/aK_E > 1$. At $K_L/aK_E = 1$ the above equation leads to a constant value for E_B.

In every experiment, the fluorimeter measures both bound ethidium E_B and free ethidium E_F from displacement by the

metal complex. Thus the asymptotic measurement of fluorescence reaches a *non zero* plateau. Mathematically, let ε represent the measured fluorescence and R_b and R_f represent the response sensitivity for the bound and free ethidium, respectively. If these are independent then it can shown that the experimental curve is given by the equation below.

$$\varepsilon = (R_b - R_f)E_B + R_f E_T \quad (21)$$

There is an asymptotic plateau given by the last term in the above equation as E_B tends to zero. Eqn (21) also highlights the fact that a differential fluorescence response is necessary for the displacement technique to be a useful measuring tool.

4 Experimental curve fitting

4.1 Non-competitive binding

To fit experimental data using a least squares technique, we employ eqn (5) directly by viewing it as a relationship between the dependent variable ε/R_b and the independent variable either S_T (for DNA titration into a fixed total metal complex concentration L_T), or, L_T (for metal complex titration into a fixed total DNA concentration S_T). To do this it is convenient to rewrite eqn (5) and (6) in terms of independent variables R, A, and B, thus,

$$y = 0.5R\left\{A + B + x - \sqrt{(A+B+x)^2 - 4Bx}\right\}. \quad (22)$$

Table 1 summarizes the meaning of the variables which are dependent on whether it is metal complex titration into the DNA or *vice versa*. The total number of DNA binding sites $S_T = B_T/n$, where B_T is the total number of bases and n the number of bases per binding site. The latter may also be estimated from the curve fitting as demonstrated in the table.

A Levenberg–Marquardt least squares routine was written to fit eqn (22) directly to experimental data and to extract optimum values for A, B and R. The code is written in C and is described by Press *et al.*[18]

As an example of the application of this method, Fig. 1 contains a set of CD spectra over the wavelength range 200 to 400 nm. The data are for the titration of ethylenediamine-1,10-phenanthrolineplatinum(II) chloride dihydrate ([Pt(en)(phen)]$Cl_2 \cdot 2H_2O$) between 0 and 7.0×10^{-5} M, into calf thymus DNA (7.62×10^{-5} bp^{-1}) in phosphate-salt (PS) buffer (1.0×10^{-3} M NaH_2PO_4, 1.0×10^{-2} M NaCl, pH 7), see Fisher.[3] The buffer and DNA background spectra have been subtracted. Fig. 2 is the experimental binding curve derived from

Table 1 Summary of the relationship between fitting and physical parameters for the metal complex and DNA titration for the case of non-competitive binding. Note: in both cases $y = \varepsilon$

Titration Variable	Fitting Parameters	Physical Parameters
$x = L_T$ B_T experimental	$A = 1/K$, $B = B_T/n$ $R = R_b$	$K = 1/A$, $n = B_T/B$ $R_b = R$
$x = B_T$ L_T experimental	$A = n/K$, $B = nL_T$ $R = R_b/n$	$K = B/AL_T$, $n = B/L_T$ $R_b = BR/L_T$

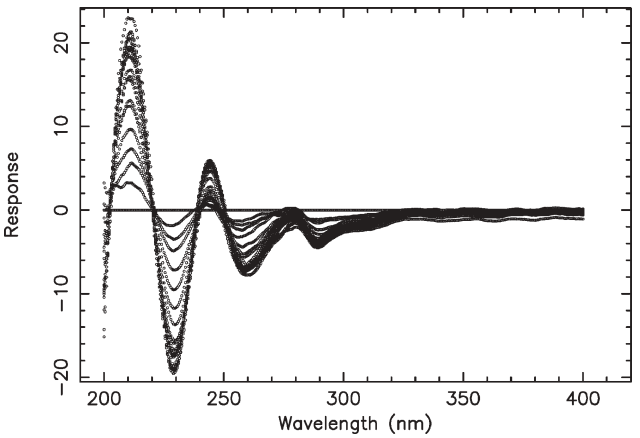


Fig. 1 CD spectra at different concentrations of metal complex [Pt(en)(phen)]$Cl_2 \cdot 2H_2O$, into calf thymus DNA in PS buffer. The data have been normalised by subtraction of the pure DNA spectrum in PS buffer.[3]

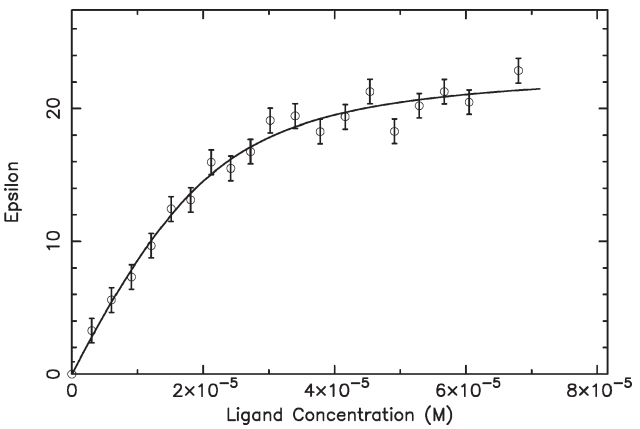


Fig. 2 Titration data of [Pt(en)(phen)]$Cl_2 \cdot 2H_2O$ derived by taking a vertical cross section at 211 nm in Fig. 1. The solid line is the theoretical fit of the data using the least squares method described in Section 4.1. Calculated parameters, averaged over 100 Monte Carlo simulations, $K = 2.6 \pm 0.1 \times 10^5$ M^{-1}, $n = 4.0 \pm 0.1$, and $R_b = 1.16 \times 10^6$.

Fig. 1 at 211 nm (the first maxima of Fig. 1). In addition this figure contains the theoretical fit generated by the least squares technique described in this paper. For negative peaks the experimental points are inverted before curve fitting.

The error bars on the experimental points have been estimated from the quality of the fit given by the minimisation of the χ^2 function,

$$\chi^2 = \sum_{i=1}^{N}\left[\frac{y_i - f(x_i,R,A,B)}{\sigma_i}\right]^2, \quad (23)$$

where y_i are the N experimental values, $f(x_i,R,A,B)$ the value of the function at x_i, and σ_i is the standard deviation of each data point. Since we do not know this, *a priori*, we can obtain an estimate of it after minimising χ^2 with respect to A, B, and R, and then making σ_i the subject of eqn (23) assuming it to be approximately the same for each data point. This σ_i value divided by $N - 3$ is the standard error displayed as the error

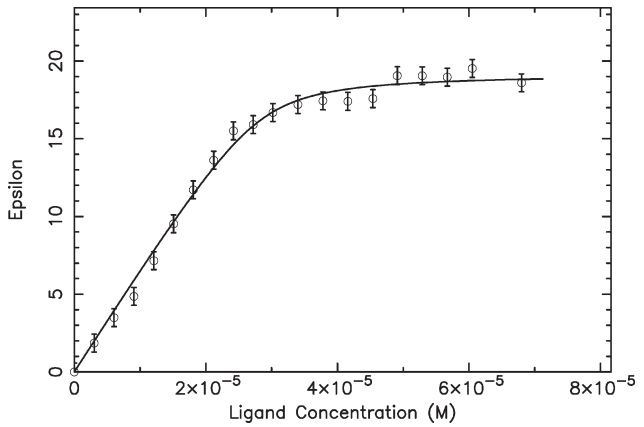


Fig. 3 Titration data at 229 nm derived from Fig. 1. Calculated parameters averaged over 100 Monte Carlo simulations: $K = 1.7 \pm 0.2 \times 10^6$ M^{-1}, $n = 2.72 \pm 0.01$, and $R_b = 6.67 \times 10^5$.

bar on the points in Fig. 2. Good fits automatically generate small error bars.

Fig. 3 is the titration curve derived from the negative minima at 229 nm in Fig. 1. The solid line is again the fit from the least squares method of this paper. For comparison, using the Intrinsic Method of Rodger and Nordén[22] on the same experimental data at 229 nm, $K = 2.12 \times 10^6$ M^{-1}, $R_b = 1.16 \times 10^{-6}$, $n = 2.9$, and $B_T = 7.62 \times 10^{-5}$ M. The latter took approximately 4 hours of operator time and involved intermediate steps with considerable judgement because of non-linearities in the Scatchard plot.

The computer based method described in this paper took less than 1 s to obtain similar values without any assumptions other than the applicability of the model described in eqn (1).

The science productivity of the least squares method together with working directly from all the data in the spectra generated by either a spectropolarimeter or spectrophotometer, is highlighted in Fig. 4, which is a plot of the equilibrium binding constant variation across the second (negative) peak in Fig. 1. It demonstrates implicit difficulties in extraction of a definitive value for the equilibrium binding

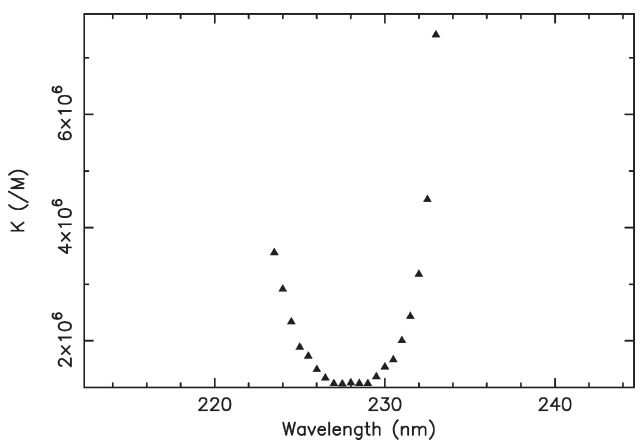


Fig. 4 Variability of the equilibrium binding constant K as a function of wavelength across the peak at 229 nm of Fig. 1 and measured between 223 nm and 233 nm.

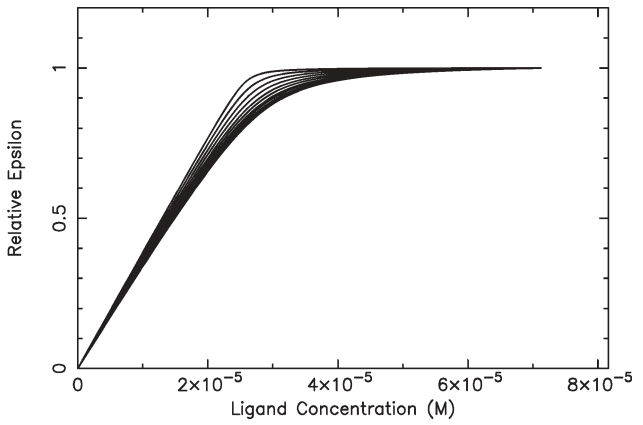


Fig. 5 Normalised theoretical curves to experimental data given in Fig. 1 between 226 nm and 234 nm. The variation in the shape of the superimposed data illustrates the measured variability of K given in Fig. 4.

constant, independent of wavelength, for metal complex DNA interaction. Fig. 3 demonstrates this again with the extracted value for K at 229 nm being almost an order of magnitude larger than that of Fig. 2 at 211 nm. Another way of demonstrating this variability is to superimpose, after normalisation, the theoretical curves fitted to experimental titration data. Fig. 5 illustrates this for wavelengths between 226 nm and 234 nm for the data of Fig. 4. The variability of K shows itself particularly in the differences in the 'knee' of the ensemble of curves. We expect that if K is constant, then all the curves should be superimposed because the response amplitudes at a given wavelength should be scalable versions of each other.

We hypothesise that this variable K phenomenon may result from the non-specific binding of the metal complex to DNA. We have applied our method to the previously reported results of Rodger *et al.*,[21] for the preferential binding of anthracene-9-carbonyl-N^1-spermine to poly [d(G-C).d(G-C)]. If a similar superimposed plot of theoretical fits to the experimental data is made over their two experimental induced CD peaks, we

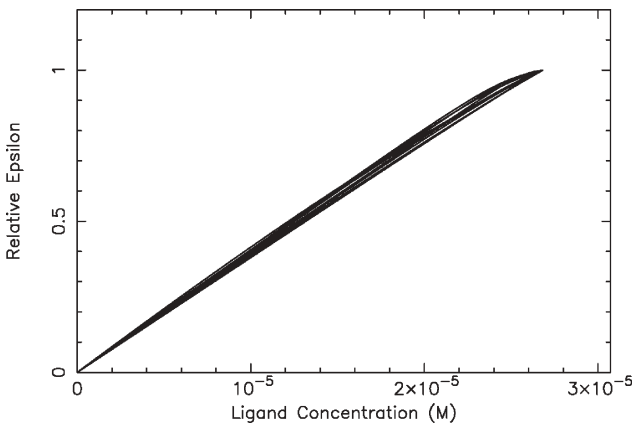


Fig. 6 Normalised theoretical curves for wavelengths 226 nm to 249 nm, derived from experimental data published by Rodger *et al.*[21] for anthracene-9-carbonyl-N^1-spermine titrated into poly [d(G-C).d(G-C)].

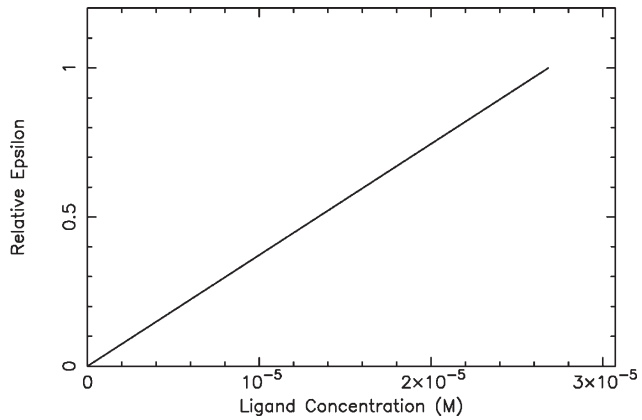


Fig. 7 Normalised theoretical curves for wavelengths 257 nm to 267 nm, derived from experimental data published by Rodger *et al.*[21] for anthracene-9-carbonyl-N^1-spermine titrated into poly [d(G-C).d(G-C)].

obtain Fig. 6 and 7. It is clear that they overlap to a much greater extent than the curves in Fig. 5, indicating that K is near constant for binding in DNA with consistent repetitive base-pair sequences. It should be noted that using our computational method we obtained a value of $K = 4.3 \times 10^7\ M^{-1}$, consistent with their published value of $K = 2.2 \times 10^7\ M^{-1}$. The difference is largely a result of the methodology. For our computational method, best estimates for K are obtained if the titrations are taken to saturation. The results of Rodger *et al.*[21] do not go to saturation. We estimate $n = 3.1$ base-pairs associated with the binding and this compares favourably with their published value of $n = 3.1$.

A Monte Carlo statistical method was devised to derive standard errors in K and n. After a least squares fit of the data and using the best estimates for A, B and R, theoretical experimental points y_i are generated as follows,

$$y_i = f(x_i,R,A,B) + G(\sigma_i), \qquad (24)$$

where $G(\sigma_i)$ is a Gaussian variable with mean of zero and standard deviation σ_i derived from eqn (23). This new set of points is fitted for best estimates of the parameters A, B and R. This Monte Carlo process is repeated as many times as necessary to obtain a desired level of standard error in these parameters. Finally K and n are calculated together with their standard errors and it is these values which are displayed in the figures. It should be noted that if σ_i in eqn (24) is set to zero that the least squares method described here retrieves the original parameters, A, B, and R with an accuracy of 1 part in 10^{-6}.

4.2 Competitive binding

As described in Section 3, displacement experiments are done in two steps. The first is an ethidium titration analysis to determine K_E and n_E, and is treated in the same manner as described in Section 4.1 for non-competitive binding. In the second step these values together with the initial conditions for E_T and $S_T = B_T/n_E$ are used to calculate α and β using eqn (18) and (19). Then for the metal complex displacement titration we write the quadratic solution to eqn (20) in a generic form in terms of a number of independent parameters: R, A, B, C and O as follows,

$$y = 0.5R\left\{A - B + Cx - \sqrt{(A - B + Cx)^2 - 4\frac{\beta}{\alpha}A}\right\} + O \qquad (25)$$

Table 2 Summary of the relationship between fitting and physical parameters for displacement titration. Note that $y = \varepsilon$ and $\gamma = (1 - K_L/aK_E)$

Titration Variable	Fitting Parameters	Physical Parameters
$x = L_T$	$A = \alpha/\gamma$, $B = S_B(1/\gamma - 1)$ $C = a(1/\gamma - 1)$, $R = R_b - R_f$ $O = R_f E_T$	$K_L = \alpha K_E C/A$ $R_b - R_f = R$ $R_f = O/E_T$ $a = C(A/\alpha - 1)$

Table 2 summarises the meaning of the variables and how they are related to physical parameters.

A second Levenberg–Marquardt least squares routine was written to fit eqn (25) directly to experimental data and to extract optimum values for R, A, B, C, and O and then from these K_L and other parameters are calculated as given in Table 2. The structure of the code is similar to that described in Section 4.1. The error analysis is also similar to that of Section 4.1 except that there are five independent parameters instead of three.

As an example of the above, the equilibrium constant K_E of ethidium bromide is measured by preparing samples containing a uniform concentration of calf thymus DNA ($7.44 \times 10^{-5}\ bp^{-1}$), in PS buffer, and a range of concentrations of ethidium bromide (0 to 8.12×10^{-5} M). Fig. 8 contains the fluorescence spectra for this titration; see Fisher[3] for comprehensive experimental detail. Fig. 9 is the ethidium bromide binding curve taken at 600 nm. The experimental data is analysed using exactly the same technique as in the non-competitive case described in Sections 2 and 4.1. As before, the methods described in this paper reveal the intrinsic dependence of the binding constant on wavelength and this is displayed in Fig. 10. We chose 600 nm as it is near the maximum height of

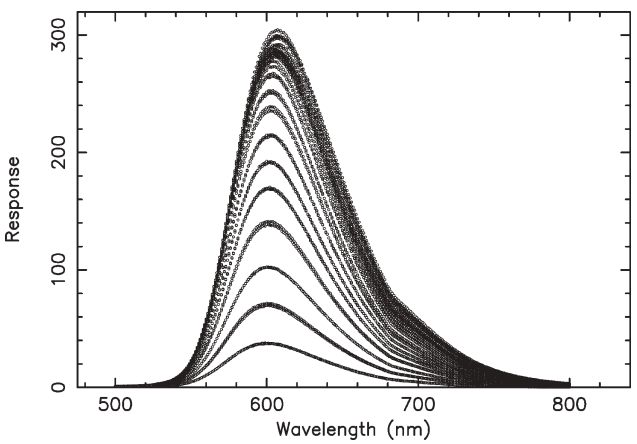


Fig. 8 Fluorescence induced displacement spectra at different concentrations of ethidium bromide, and constant calf thymus DNA in PS buffer. The base-pair DNA concentration is $7.44 \times 10^{-5}\ bp^{-1}$.

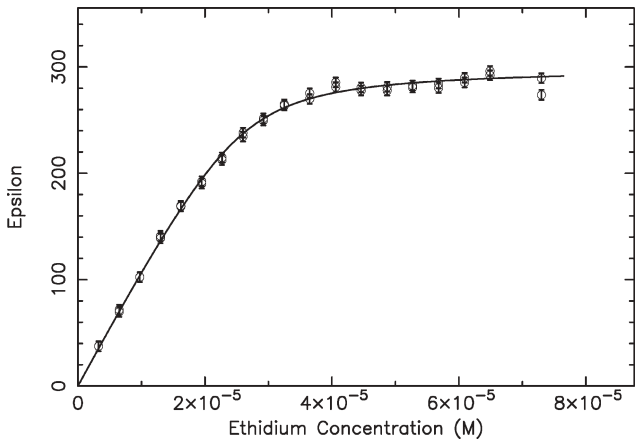


Fig. 9 Ethidium binding curve at 600 nm derived from Fig. 8. Calculated parameters, averaged over 100 Monte Carlo simulations, $K_E = 7.26 \pm 0.12 \times 10^5\ M^{-1}$, $n_E = 2.867 \pm 0.006$, and $R_b = 1.153 \times 10^7$.

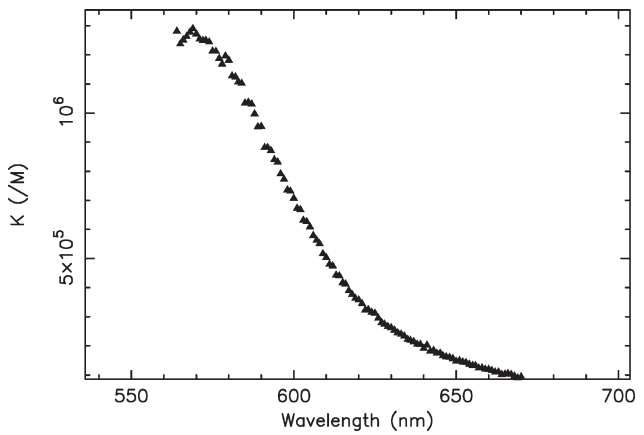


Fig. 10 Variability of the equilibrium binding constant K_E as a function of wavelength across the peak at 600 nm of Fig. 8.

the fluorescence peak as is traditionally done. At this wavelength, $K_E = 7.26 \pm 0.12 \times 10^5\ M^{-1}$, and the average number of base pairs occupied by the ethidium bromide is 2.87.

For the displacement part of the experiment, samples were prepared containing fixed concentrations of ethidium bromide (8.14×10^{-5} M) and calf thymus DNA in PS buffer (base pair concentration $7.47 \times 10^{-5}\ bp^{-1}$) and a range of concentrations of the metal complex $[Pt(en)(phen)]Cl_2 \cdot 2H_2O$ between 0 and $\sim 2.4 \times 10^{-4}$ M. We calculate α and β (eqn 18 and 19) from K_E above and $E_T = 8.14 \times 10^{-5}$ M as well as the total number of sites from the point of view of ethidium bromide to be $S_T = 7.47 \times 10^{-5}/2.87$ M. Fig. 11 displays the displacement binding curve derived from the fluorescence data at 600 nm. Also shown in this figure is the theoretical fit of eqn (25), using the methods described in this paper, and the derived parameters. An interesting aspect of this fit is that the metal complex occupies slightly more DNA base pairs than the ethidium in the ratio $n_L/n_E = 1.26$. The equilibrium binding constant for the metal complex is found to be $K_L = 9.85 \pm 0.05 \times 10^5\ M^{-1}$.

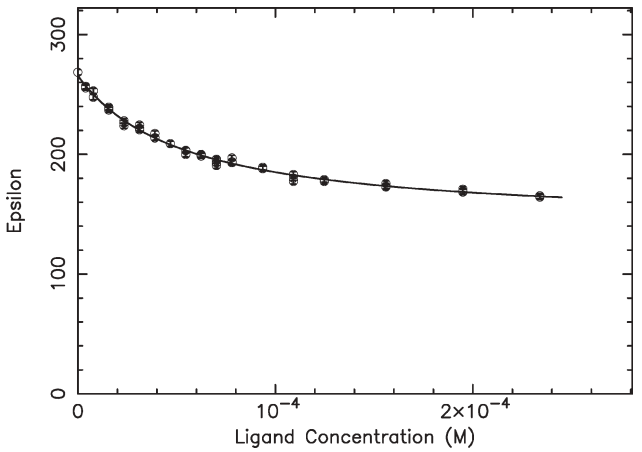


Fig. 11 Displacement of ethidium bromide by the metal complex $[Pt(en)(phen)]Cl_2 \cdot 2H_2O$. Using the method described in Section 4.2, the average values at 600 nm, over 100 Monte Carlo simulations are: $K_L = 9.85 \pm 0.05 \times 10^5\ M^{-1}$ and $n_L/n_E = 1.256 \pm 0.006$.

5 Conclusions and discussion

Presented in this paper is the underlying method for simplifying the process of data analysis to determine equilibrium binding constants from metal complex and DNA titration data for both the non-competitive and competitive (or displacement) cases. The theory derived may be fitted to experimental data using a least squares method. For the non-competitive binding case, extracted parameters are shown to be consistent with the Intrinsic Method (Rodger and Nordén[22]). The competitive or displacement theory is structurally similar, but novel, and leads to equilibrium binding constants of the metal complex displacing ethidium as well as an estimate of the relative number of bases of the DNA which are occupied. In each case the neighbour exclusion is evident in the value determined for n. Statistics for the extracted parameters is *via* a Monte Carlo technique which superimposes on the experimental data a Gaussian variation consistent with the quality of the data. An important demonstration in this paper is that binding constants are wavelength dependent.

The above mathematics forms the basis of a comprehensive computer analysis package, developed in our laboratory, which takes files directly from a fluorescence or UV-Vis spectrophotometer, or a spectropolarimeter (for CD or LD) and allows an operator to plot spectra, extract titration curves and determine equilibrium rate constants directly, without the traditional intermediate steps, and therefore, improve the speed at which the binding results can be extracted from the spectra.

As with all binding constant determinations the results are only as good as the data obtained from the experiment. There are limitations that should be noted: (i) the concentration of the biomolecule and ligand must always be within the detectable range of the spectrophotometer. (ii) In several cases we have used more than one spectroscopic tool to examine a binding where one could not adequately measure the event using a single technique.

We have routinely examined compounds with high binding affinity by fluorescence spectroscopy due to its greater

sensitivity. However, this requires that the ligand: (i) is inherently fluorescent, (ii) can displace a fluorescent molecule like ethidium bromide, or (iii) can be tagged with a fluorescent molecule. We have not taken into account the effects of fluorescent quenching[14] in this work although experiments are already underway to enable us to examine this effect. Unexpected aggregation may pose problems in any spectroscopic determination. Pulsed gradient spin-echo NMR experiments[16] of the ligand to exclude the possibility of aggregation would complement the present technique for determining the binding affinity of the ligand.

Acknowledgements

The authors wish to thank Craig Brodie from the University of Western Sydney for supplying the ethylenediamine-1,10-phenanthrolineplatinum(II) chloride dihydrate ([Pt(en)(phen)]$Cl_2 \cdot 2H_2O$).

References

1 C. R. Brodie, J. G. Collins and J. R. Aldrich-Wright, *Dalton Trans.*, 2004, **8**, 1145.
2 J. B. Chaires, *Methods Enzymol.*, 2001, **340**, 3, and references within.
3 D. Fisher, PhD thesis, the University of Sydney, Australia, 2005.
4 M. Howe-Grant, K. C. Wu, W. R. Bauer and S. J. Lippard, *Biochemistry*, 1976, **15**, 4339.
5 M. A. Ismail, K. J. Sanders, G. C. Fennell, H. C. Latham, P. Wormell and A. Rodger, *Biopolymers*, 1998, **46**, 127.
6 I. M. Klotz, *Ligand–Receptor Energetics - A Guide for the Perplexed*, Wiley Interscience, New York, 1997.
7 C. V. Kumar and E. H. Asuncion, *J. Am. Chem. Soc.*, 1993, **115**, 8547.
8 S. J. Lippard, *J. Am. Chem. Soc.*, 1978, **11**, 211.
9 H.-Q. Liu, S.-M. Peng and C. M. Che, *J. Chem. Soc., Chem. Commun.*, 1995, 509.
10 J. W. Lown, *Anthracycline and Anthracendione Based Anticancer Agents*. Elsevier, Amsterdam, 1988, ch. VII, p. 245.
11 J. D. McGhee and P. H. von Hippel, *J. Mol. Biol.*, 1974, **86**, 469.
12 Mudasir, N. Yoshioka and H. Inoue, *J. Inorg. Biochem.*, 1999, **77**, 239.
13 B. Nordén and F. Tjerneld, *Biophys. Chem.*, 1976, **4**, 191.
14 R. F. Pasternack, M. Caccam, B. Keogh, T. A. Stephenson, A. P. Williams and E. J. Gibbs, *J. Am. Chem. Soc.*, 1991, **113**, 6835.
15 D. S. Pilch, N. Poklar, E. E. Baird, P. B. Dervan and K. J. Breslaur, *Biochemistry*, 1999, **38**, 2134.
16 W. S. Price, in *Encyclopedia of Nuclear Magnetic Resonance*, ed. D. M. Grant and R. K. Harris, Wiley, New York, 2002, vol. 9, p. 364.
17 A. M. Pyle, J. P. Rehman, R. Meshoyrer, C. V. Kumar, N. J. Turro and J. K. Barton, *J. Am. Chem. Soc.*, 1989, **111**, 3051.
18 W. H. Press, S. A. Teukolsky, W. T. Vettering and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 1992.
19 X. Qu and J. B. Chaires, *Meth. Enzymol.*, 2000, **321**, 353, and references within.
20 R. Rodger, *Methods Enzymol.*, 1992, **226**, 232.
21 A. Rodger, I. S. Blagbrough, G. Adlam and M. L. Carpenter, *Biopolymers*, 1994, **34**, 1583.
22 A. Rodger and B. Nordén, *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, UK, 1997.
23 G. Scatchard, *Ann. N. Y. Acad. Sci.*, 1949, **51**, 660.
24 A. Wolfe, G. H. Shimer, Jr. and T. Meehan, *Biochemistry*, 1987, **26**, 6392.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 25 (2007) 794–800
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Shape effects on the activity of synthetic major-groove binding ligands

Syma Khalid [1], Michael J. Hannon [2], Alison Rodger, P. Mark Rodger *

Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK

Received 17 February 2006; received in revised form 5 July 2006; accepted 5 July 2006
Available online 11 July 2006

Abstract

In this work we present the results of a molecular simulation study of two different tetracationic bis iron(II) supramolecular cylinders interacting with DNA. One cylinder has been shown to bind in the major groove of DNA and to induce dramatic coiling of the DNA; the second is a derivative of the first, with additional methyl groups attached so as to give a larger cylinder-radius. The simulations show that both cylinders bind strongly to the major groove of the DNA, and induce complex structural changes in A–T rich regions. Whereas the parent cylinder tends to bind along the major groove, the derivatised cylinder tends to twist so that only one end remains within the major groove. Both G–C rich and A–T rich binding sites for the derivatised cylinder are discussed.

Keywords: DNA; Major groove; Molecular dynamics; DNA coiling; Macromolecular cylinder

1. Introduction

Advances in chemical biology over the last decade have meant that developing methods to control gene expression has become a major research goal [1–4]. One of the ways being considered for this is to use ligands that bind non-covalently to specific sequences of DNA, thereby either activating or suppressing transcription. In this, researchers are seeking to develop compounds that will mimic biological systems, where such control is often exerted through proteins that bind non-covalently and sequence-specifically in the major groove of DNA [5–7]. Recently, the first step towards this aim has been achieved with the synthesis of a dimetallic supramolecular cylinder (**1**, Fig. 1) that has been shown to bind very strongly, and non-covalently, to the major groove of DNA and to induce a very large degree of binding in the duplex [8,9]: flow linear dichroism studies indicate this cylinder bends DNA by as much as 40°–60° per bound molecule [10], whereas known bending agents such as cobalt ammines generate just 2°–5° bend per bound molecule [11].

In a recent molecular simulation study of the influence of **1** on DNA [12] the authors showed that the DNA response is surprisingly independent of the electrostatic interactions between the DNA and the cylinder.[3] In particular, a neutralised version of **1** was found to bend the DNA at least as strongly as did the parent, tetracationic, cylinder. This suggests that the biological effect of the cylinder may depend strongly on its detailed shape, as expressed through the shorter ranged van der Waals forces. It is therefore of interest to examine how sensitive the DNA/cylinder complex is to derivatisation of the cylinder. In this paper we report the results of a comparative molecular dynamics (MD) simulation study of the parent cylinder (**1**) and a hexamethylated derivative (**2**, Fig. 1) binding to the major groove of a DNA dodecamer; both cylinders were used in the M enantiomeric form. The choice of methylation site follows from our experimental and modelling deductions about the effect of this position on the interaction of the cylinders with cellulose [13].

2. Computational methods

Molecular simulations were carried out using the same protocol and potentials as were used in our previous study [12].

* Corresponding author.
E-mail address: p.m.rodger@warwick.ac.uk (P.M. Rodger).
[1] Current address: Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.
[2] Current address: School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK.

[3] Because of the overlap of coordination chemistry and biological chemistry in this paper, we shall avoid the use of the term ''ligand'', and shall instead refer to the macromolecular cylinder simply as ''cylinder''.

1093-3263/$ – see front matter
doi:10.1016/j.jmgm.2006.07.004

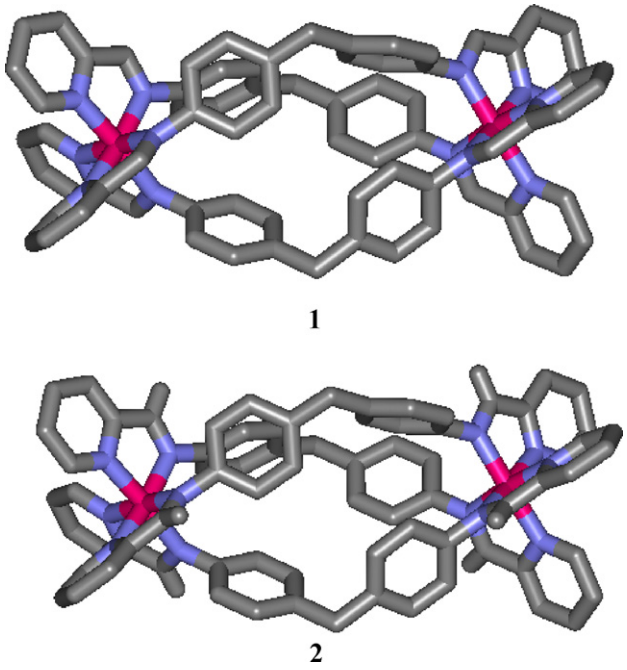


Fig. 1. The tetracationic supramolecular cylinder $[Fe_2(C_{25}H_{20}N_6)]^{4+}$ (**1**, denoted C^{4+}) and a derivative $[Fe_2(C_{31}H_{32}N_6)]^{4+}$, in which a methyl group has been added to each of the six imine carbon atoms (**2**, denoted D^{4+}). M enantiomers have been used in both cases. Colour key: Fe^{II}, pink; N, blue; C, grey; H atoms not shown.

Molecular dynamics (MD) simulations were used to simulate the response of a DNA dodecamer, d(CCCCC-TTTTTCC)•d(GGAAAAAGGGGG) to the presence of either the supramolecular cylinder **1** (C^{4+}) or its derivative **2** (D^{4+}). A four-step protocol was used to achieve this:

(1) Potential binding sites were identified from high temperature MD simulations; the DNA duplex was immobilised, the cylinder treated as a rigid body, and the environment modelled as a continuum dielectric solvent but with explicit counterions. Configurations with favourable cylinder/DNA interaction energies were identified, and the energy optimised with respect to position of the cylinder (only). The resulting structure was used as the initial DNA/cylinder/counterion conformation for subsequent calculations. Manual docking was also used to check the validity of the binding sites found. Due to the accessibility of the cylinder to the DNA grooves, comprehensive identification of binding sites was possible without resorting to more sophisticated docking methods.
(2) The DNA/cylinder dimers were solvated with explicit water molecules.
(3) A series of six short equilibration MD simulations were performed in which the DNA was harmonically restrained to the canonical B-conformation (with force constants reducing from 100 to 1 kcal/mol/Å^2). In all other respects, the full intermolecular potentials were used for these calculations.
(4) Five nanosecond MD simulations were then performed on the fully atomistic system without any restraints. As with the previous study, the major dynamical response of the DNA was complete within about 1 ns. We report here the data for the first 3 ns so as to focus on the DNA response.

A number of alternative low energy conformations were examined and all found to involve major groove binding. Several of these were then carried forward into full solvated MD simulations following the protocol given above [14]. All such simulations exhibited a very similar DNA response to that reported herein – particularly with respect to the extent to which they bent the DNA, and the resultant stability of the DNA – and confirmed that the results of our simulations were not sensitive to variations in the initial binding site. These repeat simulations also proved our results to be robust with respect to the initial background ion distribution, since each configuration involved a very different, essentially random, initial arrangement of the sodium and chloride ions.

The cylinder was modelled using the CHARMM22 force-field [15], but with the FeN_6 sub-unit treated as a rigid body with geometry taken from the crystal structure [16]. The use of a rigid FeN_6 subunit obviates the need to develop either van der Waals or intra-molecular potentials for the Fe^{II} centre, and the overall cylinder potential was found to give a good description of the crystal structure of **1** [14]. Na^+ and Cl^- were modelled with CHARMM22 and water with TIP3P [17].

We have also chosen to model the DNA duplex using CHARMM22. While it is more usual to model DNA using the CHARMM27 force-field [18,19] we note that CHARMM27 was not designed for, nor fitted to, systems with strongly interacting macromolecules such as **1**. A comparison of these two force-fields for **1** bound to DNA was reported in our earlier study [12] and revealed that the CHARMM22 force-field does have some advantages over CHARMM27 when this highly charged cationic cylinder is bound to the major groove. In particular, whereas CHARMM27 underestimated the extent of bending and produced a stable B-DNA duplex, CHARMM22 predicted that C^{4+} disrupted A–T base pairs adjacent to the cylinder binding site; similar disruption has since been confirmed in experimental crystal structures [20]. We have therefore used CHARMM22 to model the DNA interacting with **1** and **2** in this work.

Calculations were performed using DL_POLY [21]; the validity of using CHARMM force-fields within DL_POLY has been demonstrated previously [12,22]. Simulations were performed in the NVT ensemble (Hoover thermostat) with orthorhombic periodic boundary conditions and an EWALD treatment of the electrostatic interactions. Each system contained the DNA, cylinder, 3270 water molecules (TIP3P), 58 Cl^- anions and 78 Na^+ cations in a 45 Å × 45 Å × 60 Å orthorhombic simulation cell. The resulting ionic strength ([NaCl] = 0.8 M) was chosen for consistency with our previous work [12] and is similar to that used in some other MD studies [23,24]. Analysis of the structural changes was performed with 3DNA [25] and of DNA-cylinder mobility using VMD [26].

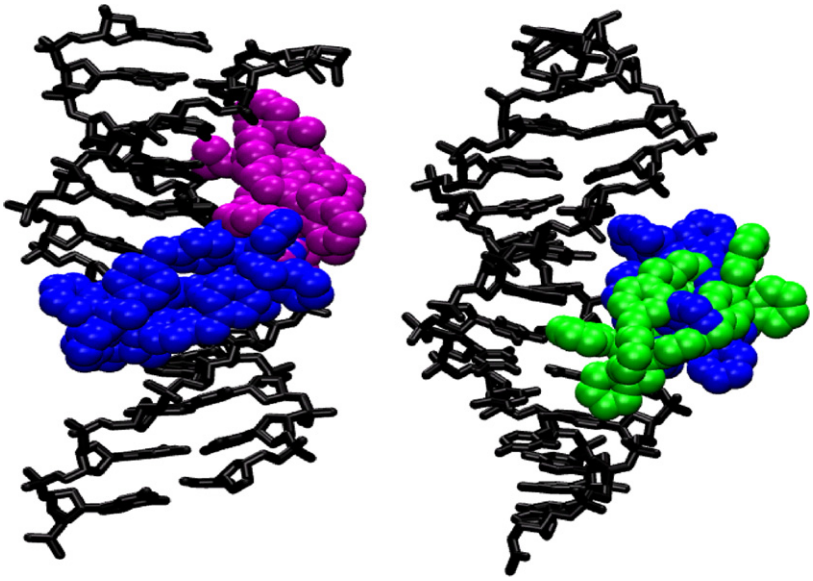


Fig. 2. DNA-cylinder starting geometries: D_{AT} (blue), D_{GC} (purple) and C_{AT} (green); the DNA is the same in the left and right figures, but is viewed from different angles to provide a clear view of the binding sites.

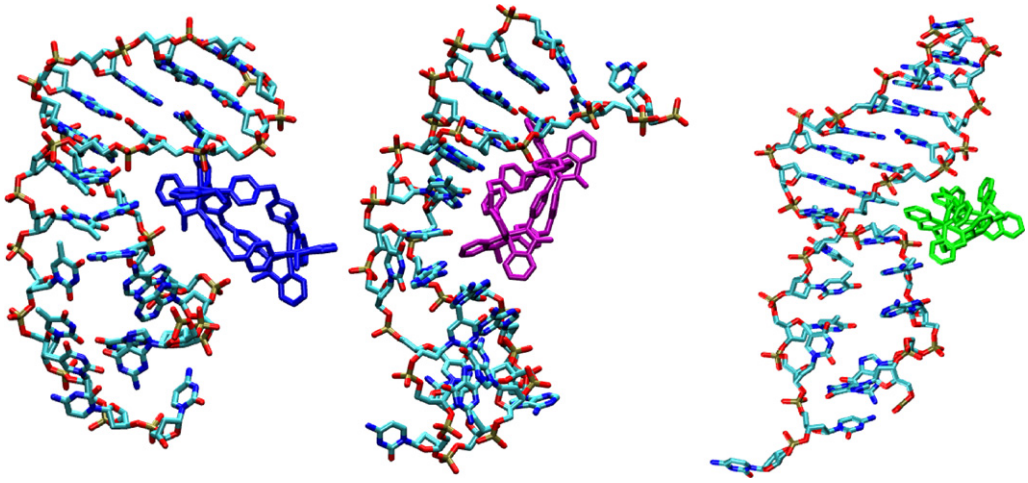


Fig. 3. DNA/cylinder configurations after 3 ns: DNA/D_{AT} (left), DNA/D_{GC} (middle) and DNA/C_{AT} (right).

3. Results

MD simulations were performed using two different low-energy major groove binding sites obtained from the docking calculations for D^{4+}: one associated predominantly with the A–T tract (base pairs 5–9) and the other with the G–C region of the DNA (base pairs 2–6). These binding sites shall be denoted D_{AT} and D_{GC}, respectively; the symbols will also be used to refer to the simulations initiated with these binding sites. The A–T binding site was found to be nearly coincident with that found [12] for C^{4+} (C_{AT}, see Fig. 2). Note that the C_{AT} binding site, in which the cylinder lies symmetrically between the two strands of the DNA spanning base pairs 6–10, is also consistent with the structure reported from NMR data [9].

Snapshots taken after 3 ns are shown in Fig. 3. A comparison of Figs. 2 and 3 provides a graphical overview of the DNA response to cylinder binding. Substantial DNA conformational changes were observed in all three cases, with curvature of the DNA backbone and disruption of the A–T base pairs, particularly below the location of the bound cylinder. After 3 ns C_{AT} remained within the major groove, but the location was no longer symmetrical with respect to the two DNA strands: there was contact[4] between the cylinder and bases C5, A16 and A17 (i.e. associated with base pairs 5, 7 and 8). The behaviour of the derivatised cylinder, D^{4+}, provides a distinct contrast. For both D_{AT} and D_{GC} the cylinder initially lay in the major groove (contact with bases A16, A17 and A18 for D_{AT}, and C2, C3, G19, G20, G21 and G22 for D_{GC}); however, it twisted during the course of the simulations so that only one end of the cylinder remained in the groove by the end of the

[4] "Contact" has been defined to occur when at least one of the inter-group atom–atom distances is less than 3.5 Å.

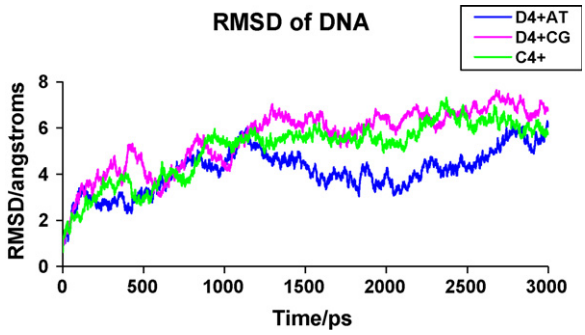


Fig. 4. Root mean square deviation of atomic positions from initial configuration; calculated using the middle 10 base pairs.

simulations. Close intermolecular distances, consistent with strong binding, were still present at the end of the simulation, with D^{4+} associated with C3 and A18 at the end of the D_{AT} simulation, but C5, A16, A17, A18, G19 and G20 for the end of the D_{GC} simulation. In all three systems there were also strong interactions between the cylinder and the backbone, but whereas C^{4+} interacted with both strands of the DNA, the D^{4+} tended to interact with just one strand by the end of the simulation (phosphate groups 16, 18 and 19 for D_{AT} and 17, 18, 19 for D_{GC}). It is intriguing to note that the two initial sites for D^{4+} resulted in very similar cylinder/DNA interactions by the end of the simulation; this is consistent with the substantial reorientation of D^{4+} already noted, and suggests that the larger size of D^{4+} compared with C^{4+} does not allow it to be constrained within specific major groove sites.

Substantial relaxation of the DNA in the presence of the cylinders is seen in each simulation, with root mean square deviations in the atomic positions in the middle ten base pairs growing to about 7 Å by 3 ns (Fig. 4). In part this was due to bending of the DNA around the cylinder, which is evident in all three systems (Fig. 3, and see discussion below), but is probably greatest and most localised with D_{AT}. However, considerable disruption of the DNA duplex was also seen in each system, particularly in the region of the A–T tract. Some indication of this is apparent in Fig. 3, but is much clearer in the analysis of the inter-nucleoside hydrogen bonding given in Table 1. Apart from some end-fraying in C_{AT}, the initial G–C tract (base pairs 1–5) is seen to be completely stable in all three systems. The A–T tract (base pairs 6–10), however, was disrupted in each simulation. This disruption was essentially complete where the cylinder was bound directly to the G–C tract, but a complex base-pair structure was observed when either cylinder was bound to the A–T tract. In these two cases, a series of base-pair mismatches were induced in the A–T rich region (indicated in bold in Table 1), which also appeared to help retain some of the original Watson-Crick base pairing (base pairs 6–19 and 7–18 for D_{AT} and 6–19, 8–17 and 10–15 for C_{AT}). Although more base pairs appear to be retained in C_{AT}, there were in fact fewer hydrogen bonds and at longer distances in this system, and it was the D_{AT} system that was more stable.

The structural changes induced in the A–T region by C_{AT} and D_{AT} was in fact both complex and dynamic, with multiple disruption, mismatch and reformation events being identified within the A–T base pairs. In the case of C_{AT}, the eighth (T8–A17), ninth (T9–A16) and tenth (T10–A15) base pairs were all seen to break and then reform over about 0.5–1.0 ns; in each case the reformed base pair remained stable for hundreds of picoseconds before subsequently being disrupted again. At least six separate mismatched pair formations (7A–17T three times, 8A–16T once and 10T–14G twice) were observed, with most of the mismatches lasting in excess of 200 ps. This dynamic behaviour was even richer for D_{AT}. For example, with

Table 1
Intra-base pair hydrogen bond distances for the average DNA calculated over the third nanosecond

Base pairings	Length/Å								
	DNA/D_{AT}			DNA / D_{GC}			DNA / C_{AT}		
	O2-N2/ O4-N6	N3-N1	N4-O6	O2-N2/ O4-N6	N3-N1	N4-O6	O2-N2/ O4-N6	N3-N1	N4-O6
[1–24] C–G	3.60	3.79	3.92	2.86	3.17	3.32			
[2–23] C–G	2.83	2.97	2.97	2.86	2.95	2.87	2.77	2.99	3.23
[3–22] C–G	2.89	2.98	2.91	2.87	2.98	2.91	2.84	2.90	2.78
[4–21] C–G	2.87	2.99	2.93	2.92	2.99	2.91	2.85	2.95	2.88
[5–20] C–G	2.88	3.01	2.96	2.81	2.90	2.84	2.94	2.99	2.88
[6–19] T–A	2.94	3.04		2.85	3.35		2.96	3.02	
[7–18] T–A	2.97	3.00					**3.53**	**2.70**	
[8–17] T–A	**3.78**						3.38		
[9–16] T–A	**3.94**						**3.34**		
[10–15] T–A	**2.32**	**3.25**					3.15		
[11–14] C–G	**3.26**								
[12–13] C–G									

The shaded boxes indicate the absence of a hydrogen bond, bold indicates hydrogen bonds between mismatched base pairs. Analysis was performed with 3DNA.

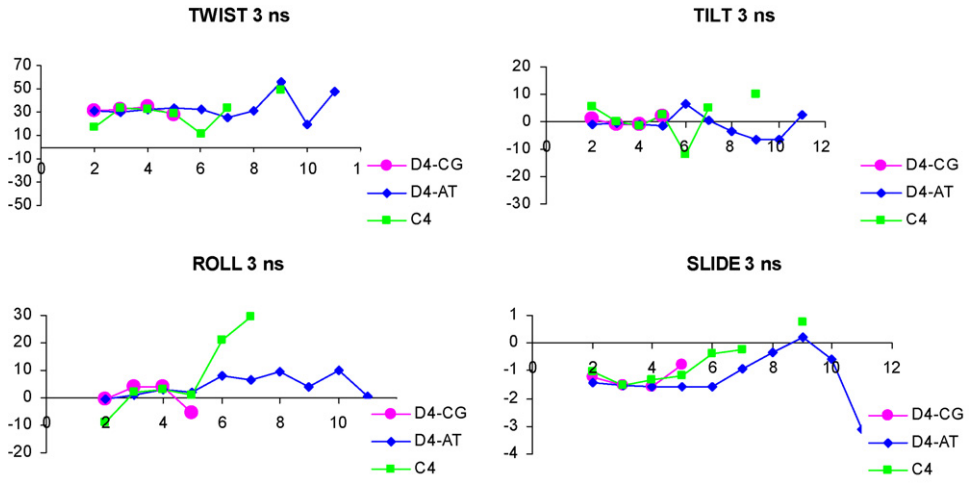


Fig. 5. Plots of the twist, tilt and roll inter-base pair angles and the slide inter-base pair distance for the DNA conformation averaged over the third nanosecond.

D_{AT} the T8–A17 and T9–A16 step was observed to form a mismatched pair (T9–A17) on four separate occasions, only to reform the two original Watson-Crick base pairs on each occasion, before a non-Watson-Crick H-bond was formed in base pair T9–A16 at about 2 ns which then remained for the rest of the simulation. It is interesting to note that none of the reformation events were evident with D_{GC}, suggesting that these cylinders are much more effective in controlling A–T base pair behaviour than the more strongly hydrogen bound G–C pairs.

It was suggested above that the dramatic conformational changes in the DNA evident in Fig. 3 were not limited to the disruption of the base pairs, but also include bending of the DNA backbone. However, it is difficult to be definitive about this. The analysis of DNA curvature is usually performed in terms of the relative geometry of adjacent base pairs (slide distance and tilt, roll and twist angles), but such an analysis presupposes the presence of stable Watson-Crick base pairs within the duplex. It is not clear to what extent such an analysis is still meaningful in the presence of the deformations in the base pairings in the A–T tract noted above. Never-the-less, base pair geometrical parameters, as calculated with 3DNA, are presented in Fig. 5 for the three DNA/cylinder systems. Parameters have been calculated for the average DNA structure observed during the third nanosecond of each simulation. Given the stable base-pairings apparent in Table 1, we expect parameters for the first five steps to be reliable and unambiguous, although interpretations based on later steps (6–11) should be treated with some caution. We note that for the first five steps (base pairs C1–G24 through T6–A19), the geometric parameters show almost no variation between the first and third nanoseconds, and so only data for the third nanosecond are presented here.

The step parameters indicate that the DNA adopts a structure that is not consistent with the canonical form of either A or B-DNA. Slide values of less than −1 Å, usually indicative of A-DNA, were observed for the D_{AT} and D_{GC} simulations and values intermediate between those expected for A and B DNA

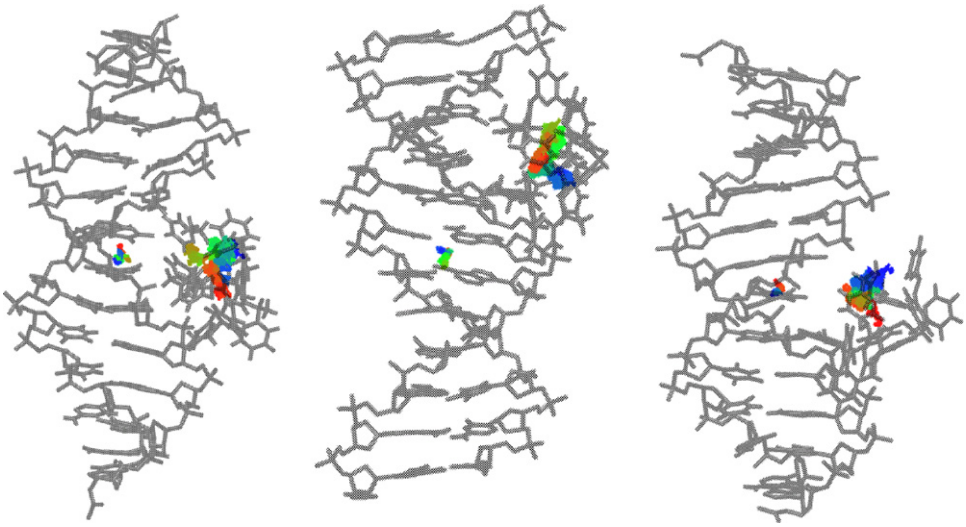


Fig. 6. D_{AT} (left), D_{GC} (middle) and C_{AT} (right). A RGB colour scale is used to represent the time dependence of the centre of mass positions (red is at the beginning of the simulation, blue is at the end).

were found in C_{AT}. On the other hand, whereas the A form of DNA is usually associated with negative values for the roll angle, all three cylinders induced positive roll angles, as is more often associated with the B-form. The twist angles in all three simulations were intermediate between those expected for the A and B forms of DNA. We conclude that these cylinders do have a substantial effect on the DNA conformation, but that the nature of this effect is difficult to categorise in terms of canonical forms—particularly with the relatively short DNA strands (dodecamers) considered here. Simulations with longer segments of DNA are in progress.

Some interesting differences were also observed in the relative motion of the DNA and cylinder. Visual inspection of the three trajectories revealed substantial movement of the cylinder with respect to the DNA. As commented earlier, the C_{AT} cylinder appears to remain within the major groove throughout the simulations, although one end is in closer contact with the DNA than the other by the end of the simulation, while the larger D^{4+} twists much more prominently so that the cylinder lies only partially within the groove for much of the D_{AT} and D_{GC} trajectories.

We have used VMD [26] to monitor the time dependence of the centres of mass of the cylinder and the DNA, and also the displacement vector between the two centres of mass. Plots of the centre of mass trajectories are given in Fig. 6. In each of the three simulations there is a much greater spread in the distribution of the cylinder centre of mass compared with the DNA centre of mass. A convenient measure of this spread is to calculate the root mean square distance (RMSD) of the centre of mass from the centre (i.e. average) of its own trajectory. Values for this (averaged over the first 3 ns) are listed in Table 2. The RMSD for the cylinder is a factor of 3.2–3.4 larger than that for the DNA duplex. The mean square displacement can be expressed as a double integral of the velocity time correlation function, and so should show a linear dependence on kT/m, where k is Boltzmann's constant, T the temperature and m is the molecular mass; thus the RMSD should scale with the square root of the molecular mass. The DNA dodecamer is 6.7 times heavier than the cylinder, which would correspond to a factor of 2.6 in the RMSDs. While these two factors are comparable, the observed value is significantly larger and so indicates some correlation in the motion—possibly because the rotation of the DNA is being superimposed on the cylinder motion through its binding to the major groove.

Table 2
Root mean square distance along a trajectory

System	Root mean square distance from the origin (Å)		
	DNA	Cylinder	Displacement
C_{AT}	0.40	1.36	1.75
D_{AT}	0.70	2.26	2.96
D_{GC}	0.59	1.91	2.49

The origin is taken as the geometric centre of the trajectory (defined by the time average x, y and z coordinates). This "displacement" refers to the vector from the DNA centre of mass to the cylinder centre of mass, and is again shifted so that the time average of this vector is taken as the origin.

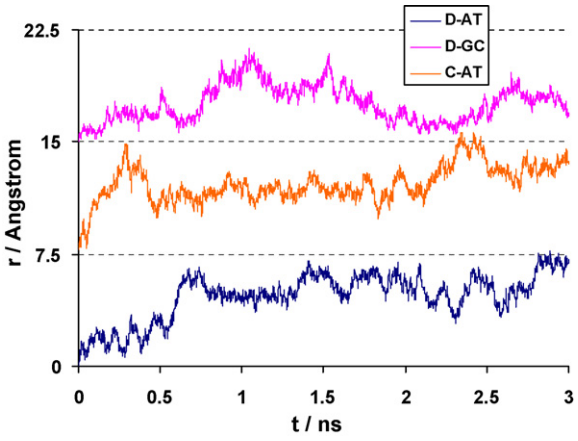


Fig. 7. Root mean square distance along a trajectory, as measured from the initial position. The trajectory is for the displacement vector from the DNA centre of mass to the cylinder (cylinder) centre of mass. Curves have been displaced vertically by 7.5 Å (C_{AT}) and 15 Å (D_{GC}) for clarity.

The motion of the cylinder relative to the DNA can be monitored in more detail by considering the distance:

$$d(t) = |\boldsymbol{r}_{DNA\rightarrow Cyl}(t) - \boldsymbol{r}_{DNA\rightarrow Cyl}(0)|$$

where $\boldsymbol{r}_{DNA\rightarrow Cyl}(t)$ is the vector from the DNA centre of mass to the cylinder centre of mass, both at time t. Plots of this quantity for the three systems are given in Fig. 7. All three systems show maximum distances of no more than 8 Å which, given the cylinders have molecular lengths of ca. 20 Å, may be explained by the reorientation of the cylinder so that it tends to point out from the major groove rather than lie within it. For both binding sites, the D^{4+} cylinder shows an initial stable region, with little relative motion during the first 0.5 ns, followed by a relatively rapid rearrangement to a new position about 5 Å removed from its initial location. In contrast, the position of the C^{4+} cylinder relaxes over the first 0.5 ns, albeit again finding a stable location about 5 Å from its initial position. Most intriguingly, the initial behaviour of D_{AT} is oscillatory, with a period of about 100 ps, suggesting that while the initial binding site was only metastable, there was a substantial energy barrier to reorientation of the cylinder that was associated with relatively low-frequency collective motions in the DNA. Similar oscillations were not apparent with the smaller C^{4+}, which fitted into the major groove more easily; nor were they apparent when the binding was within the G–C tract. It is interesting to speculate whether such coupling to vibrational motion in the DNA may be the major-groove analogue of similar couplings seen for synergistic binding of Hoechst 33258 in the minor groove [27].

4. Conclusions

The simulations presented in this article reveal a rich and complex interaction between DNA and a new class of major groove binding ligands. These DNA-binding ligands are dimetallic macromolecular cylinders, such as **1** and **2**, and have dimensions (ca. 20 Å length and 10–12 Å radius) designed to match the major groove of DNA. As anticipated from our earlier work [12], the nature of the DNA response is sensitive to

the shape of the cylinder. In this work we have shown that a hexamethylated cylinder (D^{4+}, **2**), which has a significantly larger cylindrical radius than the parent compound (C^{4+}, **1**), binds to DNA in a mode that is not completely contained within the major groove. Instead, the cylinder reorients so that only one end occupies the major groove. Never-the-less, this larger cylinder still has a very strong effect on the conformation and stability of the DNA. Depending on the initial location of the binding site, the D^{4+} may either disrupt A–T base pairs (when bound adjacent to them, but within a region rich in G–C pairs), or induce a complex set of base-pair mismatches when bound within the A–T-rich region. These mismatches were similar to those already reported for the parent cylinder, C^{4+} [12], even though the binding geometries proved to be different with these two cylinders. We conclude that the dimetallic macromolecular cylinders do induce a response in the DNA that is dependent on the DNA sequence, and that the nature of this response can be varied by altering the shape of the cylinder.

Acknowledgements

This work was supported by the EPSRC (S.K.; postgraduate studentship) and the European Union (MARCY RTN; HPRN-CT-2002-00175).

References

[1] See, for example, J.M. Berg, J.L. Tymoczko, L. Stryer, Biochemistry, 5th ed., Freeman, New York, 2002.
[2] J.P. Baak, F.R. Path, M.A. Hermsen, G. Meijer, J. Schmidt, E.A. Janssen, Eur. J. Cancer 39 (2003) 1199–1215.
[3] C.F. Calkhoven, C. Muller, A. Leutz, Trends Mol. Med. 8 (2002) 577–583.
[4] L. Rogge, Arthritis Res. Ther. 5 (2003) 47–50.
[5] R.R. Sinden, DNA Structure and Function, Academic, London, 1994.
[6] R.E. Dickerson, Nucleic Acids Res. 26 (1998) 1906–1926.
[7] C. Branden, J. Tooze, Introduction to Protein Structure, 2nd ed., Garland, New York, 1999.
[8] I. Meistermann, A. Rodger, V. Moreno, M.J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P.M. Rodger, J. Peberdy, C.J. Isaac, M.J. Hannon, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 5069–5074.
[9] E. Molderheim, I. Meistermann, A. Rodger, M.J. Hannon, E. Sletten, J. Biol. Inorg. Chem. 7 (2002) 770–780.
[10] M.J. Hannon, V. Moreno, M.J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C.J. Isaac, K.J. Sanders, A. Rodger, Angew. Chem. Int. Ed. Engl. 40 (2001) 879–884.
[11] A. Rodger, K.J. Sanders, M.J. Hannon, I. Meistermann, A. Parkinson, D.S. Vidler, I.S. Haworth, Chirality 12 (2000) 221–236.
[12] S. Khalid, M.J. Hannon, A. Rodger, P.M. Rodger, Chem. Eur. J. 12 (2006) 3493–3506.
[13] S. Khalid, P.M. Rodger, A. Rodger, J. Liq. Chromatogr. Rel. Technol. 28 (2005) 2995–3003.
[14] S. Khalid, PhD Thesis, University of Warwick, 2004.
[15] A.D. MacKerell Jr., D. Bashford, M. Bellot, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnik, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586–3616.
[16] M.J. Hannon, C.L. Painting, J. Hamblin, A. Jackson, W. Errington, Chem. Commun. (1997) 1807–1808;
C.L. Painting, PhD Thesis, University of Warwick, 1999; I. Meistermann, L.J. Childs, F. Tuna, C.L. Painting, W. Errington, N.W. Alcock, A. Rodger, M.J. Hannon, in preparation.
[17] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926–935.
[18] N. Foloppe, A.D. MacKerell Jr., J. Comp. Chem. 21 (2000) 86–104.
[19] A.D. MacKerell, N. Banavali, N. Foloppe, Biopolymers 56 (2000) 257–265.
[20] A. Oleksi, A.G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M.J. Hannon, M. Coll, Angew. Chem. Int. Ed. Engl. 45 (2006) 1227–1231.
[21] W. Smith, C.W. Yong, P.M. Rodger, Mol. Sim. 28 (2002) 385–471.
[22] M.A.S. Miguel, R. Marrington, P.M. Rodger, A. Rodger, C. Robinson, Eur. J. Biochem. 270 (2003) 3345–3352;
M.A.S. Miguel, P.M. Rodger, Phys. Chem. Chem. Phys. 5 (2003) 575–581;
R. Lukac, A.J. Clark, M.A.S. Miguel, A. Rodger, P.M. Rodger, J. Mol. Liq. 101 (2002) 261–272;
T. Astley, G.G. Birch, M.G.B. Drew, P.M. Rodger, J. Phys. Chem. A 103 (1999) 5080–5090.
[23] M. Feig, B.M. Pettitt, Biophys. J. 75 (1998) 134–149.
[24] N. Korolev, A.P. Lyubartsev, A. Laaksonen, L. Nordenskiöld, Biophys. J. 82 (2002) 2860–2875.
[25] X.J. Lu, W.K. Olson, Nucleic Acids Res. 31 (2003) 5108–5121.
[26] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 14 (1996) 33–38.
[27] S.A. Harris, E. Gavathiotis, M.S. Searle, M. Orozco, C.A. Laughton, J. Am. Chem. Soc. 123 (2001) 12658–12663.

Contributions of Hydroxyethyl Groups to the DNA Binding Affinities of Anthracene Probes

Michael R. Duff,[†] Willy B. Tan,[†] Akhilesh Bhambhani,[†] B. Scott Perrin, Jr.,[†] Jyotsna Thota,[†] Alison Rodger,[‡] and Challa V. Kumar*[,†]

Department of Chemistry, University of Connecticut, Storrs, Connecticut 06269-3060, and Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK

Received: June 26, 2006

Contributions of hydroxyethyl functions to the DNA binding affinities of substituted anthracenes are evaluated by calorimetry and spectroscopy. Isothermal titration calorimetry indicated that binding of the ligands to calf thymus DNA (5 mM Tris buffer, 50 mM NaCl, pH 7.2, 25 °C) is exothermic. The binding constants increased from 1.5×10^4 to 1.7×10^6 M^{-1} as a function of increase in the number of hydroxyethyl functions (0−4). DNA binding was accompanied by red-shifted absorption (~630 cm^{-1}), strong hypochromism (>65%), positive induced-circular dichroism bands, and negative linear dichroism signals. DNA binding, in general, increased the helix stabilities to a significant extent ($\Delta T_m \approx 7$ °C, $\Delta\Delta H \approx 3$ kcal/mol, $\Delta\Delta S \approx 6-20$ cal/K·mol). The binding constants showed a strong correlation with the number of hydroxyethyl groups present on the anthracene ring system. Analysis of the binding data using the hydrophobicity parameter (Log *P*) showed a poor correlation between the binding affinity and hydrophobicity. This observation was also supported by a comparison of the affinities of probes carrying *N*-ethyl ($K_b = 0.8 \times 10^5$ M^{-1}) versus *N*-hydroxyethyl side chains ($K_b = 5.5 \times 10^5$ M^{-1}). These are the very first examples of a strong quantitative correlation between the DNA binding affinity of a probe and the number of hydroxyethyl groups present on the probe. These quantitative findings are useful in the rational design of new ligands for high-affinity binding to DNA.

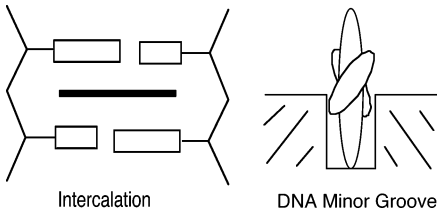
Introduction

A better understanding of DNA−ligand interactions is essential for the rational design of ligands to influence key biological processes such as gene expression, transcription, or translation.[1] Although hydrogen-bonding groups are known to play a key role in these interactions,[2] quantitative data on their contributions are still sparse.[3] One approach is to systematically increase the number of potential H-bonding side chains on a ligand and quantify their contributions to the DNA binding affinities.[4] Focus of the current investigations is to quantify the interactions of hydroxyethyl side chains by calorimetric and spectroscopic studies.

Anthracene derivatives are chosen, in the current investigations, for the following reasons. Several anthracene derivatives are known to have significant biological activities against specific types of cancers.[5] Pseudourea, for example, was one of the earliest examples of anthracene-based drugs tested in clinical trials,[6] and anthracene itself was reported to be effective against specific skin ailments.[7] The planar, linear, three-ring system of the anthracene nucleus has potential for significant overlap with the DNA base pairs, and it is expected to facilitate intercalation of the probe into the DNA helix (Chart 1).[8] The versatile chemistry of anthracene nucleus provides a convenient route to prepare a number of closely related derivatives.[9]

Additional advantages of the anthracene probes are that: (1) they have moderate absorption cross-sections in the near-UV region and good fluorescence quantum yields, which are useful to monitor ligand binding to DNA by spectroscopic methods;[10] (2) while the GC sequences of DNA quench the fluorescence of most anthracene derivatives, AT sequences enhance anthryl fluorescence, and this provides a useful marker to identify the nature of the binding site;[11] (3) anthryl probes have long-lived triplet excited states, which can be used to induce significant DNA damage and strand cleavage;[12,13] (4) substituents at the 9 and 10 positions of the anthracene nucleus are strategically positioned such that these occupy the grooves when the anthracene moiety is intercalated into the helix;[14] and (5) the spectroscopic signatures of the DNA binding modes of anthracene probes have been recently identified,[15] and these will be useful in the analysis of the current data.

CHART 1: Two Major Binding Modes (Intercalation and Groove Binding) of Anthryl Ligands to the DNA Double Helix[a]



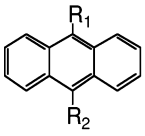
[a] Only one geometry for the minor groove binding is shown, and the small ovals in the front and the back of the large oval represent the 9,10-substituents on the anthracene ring.

The number of hydroxyethyl substituents on the anthracene nucleus is increased from 0 to 4 by substitution at the 9 and 10 positions (**1**−**4**, Chart 2). The hydroxyethyl side chain is chosen for these studies due to its ability to support potential hydrogen-bonding contacts with the helix. These functions have been implicated in enhanced binding affinities of some metal complexes and anticancer agents.[5b,16] However, there are no clear, quantitative correlations between the DNA binding affinities and the number of such functions present on the ligand.

* Corresponding author. E-mail: challa.kumar@uconn.edu.
[†] University of Connecticut.
[‡] University of Warwick.

10.1021/jp063997m CCC: $33.50
Published on Web 09/22/2006

CHART 2: Structures of the Anthracene Derivatives Discussed in This Study



#		R_1	R_2
1.	AMAC	$CH_2NH_3^+$	H
2.	MEA	$CH_2NH_2^+CH_2CH_2OH$	H
3.	BMEA	$CH_2NH_2^+CH_2CH_2OH$	$CH_2NH_2^+CH_2CH_2OH$
4.	BDEA	$CH_2NH^+(CH_2CH_2OH)_2$	$CH_2NH^+(CH_2CH_2OH)_2$
5.	APAC	$CH_2CH_2CH_2NH_3^+$	H
6.	N-Et-AMAC	$CH_2NH_2^+CH_2CH_3$	H

The hydroxyethyl group is similar in size, flexibility, conformation, and polarity to the propylamine side chain of **5**, and the latter was suggested to make H-bonding contacts with the DNA minor groove.[14] However, the OH groups of **2**−**4** are positioned one covalent bond farther away from the anthryl ring than the NH_2 group of **5**, and this feature is expected to allow higher flexibility for the side chain.

Among the probes shown in Chart 2, probes **1** and **4** indicated enhanced binding affinity with **4** when compared to that of **1**.[10,14,15] However, it is not clear if this increased affinity is due to the increased number of hydroxyethyl functions or due to increased hydrophobicity of the probe. Anthracene derivatives **2** and **3** were prepared, therefore, and they have been examined to evaluate the role of the side chains in the binding interaction. By comparing the binding properties of **1**−**4**, it is possible to evaluate the contributions of these side chains to the overall binding affinity. Comparison of the binding properties of **2** with those of **6** would reveal the contributions of the OH group to the binding interactions. Quantitative data on the binding of these probes to calf thymus DNA are presented here, and these indicate a strong correlation between the number of hydroxyethyl functions present on the probe and their binding affinities. Such correlations are useful for the rational design of ligands of high DNA binding affinities.

Experimental Section

DNA Samples. Calf thymus DNA (CT DNA, type I) was purchased from Sigma Chemical Co. (St. Louis, MO) and purified according to published protocols.[17] All experiments were carried out in Tris buffer (5 mM Tris, 50 mM NaCl, pH adjusted to 7.2), unless mentioned otherwise. Stock solutions of CT DNA were prepared by dissolving an appropriate mass of DNA in Tris buffer. All DNA concentrations are expressed in terms of base pairs of DNA.

Synthesis of 9-(*N*-(2-Hydroxyethyl)aminomethyl)anthracene (MEA, 2). A solution of 9-chloromethylanthracene (50 mg, 0.22 mmol) in THF (15 mL) was added to a solution of ethanolamine (61 mg, 1 equiv, 10 mL of THF) over a period of 0.5 h. The reaction was allowed to proceed for 24 h, and progress was checked by monitoring the absorption spectrum of the reaction mixture. Disappearance of the 397 nm absorption peak due to the 9-chloromethylanthracene was noted, which was accompanied by the gradual growth of the product band at 388 nm. The reaction mixture was filtered, concentrated under vacuum, and allowed to crystallize. The crude product was further purified by crystallization from methanol, in 54% yield, mp 116−119 °C, and the melting point matched with the reported value.[18] ^{1}H NMR (CH_3OD) ppm: 8.43 (s, 1), 8.35 (d, 2), 8.00 (d, 2), 7.52 (t, 2), 7.44 (t, 2), 4.71 (s, 3), 3.69 (m, 2), 2.93 (m, 2). Mass spectrum indicated a molecular ion peak at (*m*/*z*) 252.

Synthesis of 9,10-Bis(*N*,*N*-(2-hydroxyethyl)aminomethyl)-anthracene (BMEA, 3). A solution of 9,10-bisbromomethylanthracene in THF was added to excess ethanolamine (2.1 equiv in THF). Reaction with the bromo-derivative was faster than with the chloro-derivative described above, and the former resulted in cleaner reactions. The product was crystallized from methanol in 23% yield, mp 157−160 °C, and the melting point matched with the reported value.[19] ^{1}H NMR (CH_3OD) ppm: 8.45 (m, 4), 7.65 (m, 4), 5.03 (s, 2), 3.81 (t, 4), 3.14 (t, 4). Mass spectrum indicated a molecular ion peak at (*m*/*z*) 325.

Even though it is appropriate to refer to these ligands by the numbers shown in Chart 2, it is more convenient to use the corresponding acronyms, and these are used from this point on in this Article.

Isothermal Titration Calorimetry (ITC). ITC studies were performed on a MicroCal VP-ITC microcalorimeter (MicroCal Inc., Northampton, MA). Anthryl derivative (∼350−170 μM) was titrated into CT DNA (100 μM DNA, Tris buffer). Multiple injections of 8−16 μL of ligand solution into a solution of CT DNA (1.4167 mL) were carried out at specific intervals, until the binding was saturated. Samples were allowed to equilibrate between injections. The entire experiment was controlled by software supplied by MicroCal, Inc., and the data were fit using the Origin 5.0 software package. Multiple data sets were acquired and analyzed using eq 1. Analysis with different starting values resulted in a consistent set of parameters, which are within our experimental error.

$$Q = 0.5(nM_t\Delta HV_o)[1 + (X_t/nM_t) + (1/nK_bM_t) - \sqrt{[(1 + (X_t/nM_t) + (1/nK_bM_t))^2 - 4X_t/nM_t]}] \quad (1)$$

The heat released (Q) during the titration is related to the molar heat of ligand binding (ΔH), volume of the sample cell (V_o), the bulk concentration of the ligand (X_t), concentration of DNA (M_t), the intrinsic binding constant (K_b), and the number of binding sites (n) by eq 1. The above equation was derived from basic thermodynamic principles, under equilibrium conditions.[20] Thermodynamic parameters for the binding such as ΔH, ΔS, binding constant (K_b), and the binding site size (n) are calculated directly from the ITC data.

Absorption Studies. Absorption spectra were recorded on a diode array HP 5750 spectrophotometer. A solution of the anthryl probe (11 μM MEA or 12 μM BMEA, 1 mL, Tris buffer, in a 1-cm path length cuvette) was titrated by adding 5 μL increments of CT DNA (3−6 mM, Tris buffer), while monitoring the absorbance in 300−500 nm region. The total volume of CT DNA added during the titration did not exceed 10% of the starting volume, and the data have been corrected for volume changes. The absorption data were analyzed using Scatchard eq 2 to estimate the binding constants.

$$1/C_f = K_bn(1/r) - K_b \quad (2)$$

In eq 2, C_f is the concentration of the free probe, n is the binding density, and r is the ratio of concentration of bound probe (C_b) to that of DNA. The absorbance data are used to calculate C_f by using the relations $C_f = (A_{obs} - \epsilon_bC_t)/(\epsilon_f - \epsilon_b)$, where A_{obs} is the absorbance at the peak position of the probe−DNA mixture, C_t is the total concentration of the probe, while ϵ_f and ϵ_b are the extinction coefficients of the free and bound probes, respectively. The value of ϵ_b was estimated from a plot of A_{obs} versus 1/[DNA]. In cases where A_{obs} versus 1/[DNA] plot was not linear, the absorption spectrum of the probe was obtained from a mixture of probe and a high concentration of CT DNA (1 mM). The intrinsic binding constant (K_b) was estimated from these data by plotting $1/C_f$ versus $1/r$ and by fitting the data to eq 1. Alternatively, the data were also analyzed using the McGhee and von Hippel eq 3 for comparison.[21]

$$r/c_f = K_b(1 - nr)[(1 - nr)/[1 - (n - 1)r]]^{n-1} \quad (3)$$

The percent hypochromism was determined from the estimated values of ϵ_f and ϵ_b. The free energy of the binding ($\Delta G°$) was calculated from eq 4.

$$\Delta G° = -RT \ln K_b \quad (4)$$

Induced Circular Dichroism (CD) Studies. The CD spectra were recorded on a Jasco J-710 spectropolarimeter (Jasco Inc., Easton, MD) interfaced with an IBM-PC computer. A mixture of the probe (75 μM) and CT DNA (300 μM) solutions was placed in a quartz cuvette of 1-cm path length, for recording the spectra in the 325−425 nm window. A cuvette of 2-mm path length was used for recording the CD spectra in the 200−325 nm window. Parameters of 1 nm bandwidth, 20 mdeg sensitivity, and 4 s response time were used to record the spectra.

Linear Dichroism Studies. Flow linear dichroism (LD) spectra were recorded on a Jasco J-715 spectropolarimeter adapted for LD measurements by using a cylindrical flow cell. The flow is generated by the rotation of an inner quartz cylinder, spaced 0.5 mm from a stationary steel/quartz cylinder.[22] Light from the source passes through the center of the cell and through the solution. The ratio of the concentration of DNA to that of the probe was varied at each stage by adding ligand solution, as well as an appropriate volume of DNA solution, so that the DNA concentration (and hence the sample viscosity) is constant. All LD spectra were measured at pH = 6.9 (1 mM sodium cacodylate, 50 mM NaCl).

DNA Melting Experiments. DNA melting studies were carried out on a Hewlett-Packard 8452A diode array spectrophotometer equipped with an HP 89090A peltier thermostat, controlled by an HP Vectra 386/16N computer. The helix melting temperatures were measured by monitoring the sample absorbance at 270 nm (Tris buffer).

Differential Scanning Calorimetry. The DSC experiments were performed on a model 6100 Nano-II differential scanning calorimeter from Calorimetry Sciences Corporation (Lindon, UT). The calorimeter had a cell volume of 0.299 mL, and it was interfaced with a Dell personal computer and controlled by software supplied by the manufacturer. In a series of DSC scans, both of the cells were first loaded with Tris buffer (5 mM Tris, 25 mM NaCl, pH 7.2), equilibrated at 10 °C for 10 min, and then the temperature was raised from 10 to 100 °C at a scan rate of 2 °C/min. The buffer versus buffer scan was repeated once more, and, after cooling to room temperature, the sample cell was emptied, rinsed, and loaded with the sample.

The samples and reference solutions were degassed for at least 5 min at room temperature and carefully loaded into the cells to avoid bubbles. During the scans, constant pressure of 3 atm was maintained over the solutions to prevent possible degassing of the samples during heating. A background scan, recorded with the buffer in both cells, was subtracted from each test scan. A typical example consisted of a scan of CT DNA (60 μM, Tris buffer) in the presence or absence of the probe (33 μM) from 10 to 100 °C, at a heating rate of 2 °C/min.

The excess molar heat capacity (C_p) was calculated from the raw data, using the mean molecular mass of 330 g/mol of nucleotides, and the partial specific volume of CT DNA has been calculated to be 0.54 mL/g.[23] Thermodynamic parameters of the helix denaturation were calculated, and reported values are averages of at least three measurements.

Sonication of CT DNA. Sonicated CT DNA was prepared by following a reported procedure.[24] A 50 W Cole-Parmer ultrasonic processor equipped with a 3 mm titanium microtip was used to sonicate the DNA sample. CT DNA (40 mL, ∼1 mg/mL, Tris buffer) was placed in a plastic test tube, cooled in an ice bath, and nitrogen gas bubbled through the DNA solution for 15 min. The solution was sonicated for 5 min followed by purging with nitrogen for 5 min. This cycle was repeated 30 times, and, after the 30th time, the solution was purged with nitrogen for 15 min. The DNA solution was again sonicated for a total of 330 s (10 s sonication followed by 5 min purging with nitrogen), and the solution was dialyzed five times against Tris buffer. The average molecular weight of DNA was estimated by measuring viscosity, using eq 5.[24] The intrinsic viscosity, [η], was 1.96 dL g^{-1}, and the average molecular weight of sonicated DNA has been estimated to be 3.0×10^5 Da. The sonicated DNA sample had a A_{260nm}/A_{280nm} ratio of 1.83, and the melting temperature of sonicated CT DNA matched with that of the unsonicated sample. Hyperchromicity (H, eq 6, where $A_{25°C}$ and $A_{96°C}$ are absorbances at 25 and 96 °C, respectively) of the sonicated DNA solution was comparable to that of the unsonicated sample.

$$[\eta] = 1.45 \times 10^{-6}\, M^{1.12} \quad (5)$$

$$H = (A_{96°C} - A_{25°C})/A_{25°C} \quad (6)$$

Computational Modeling. Selected properties of the anthracene probes were calculated by Computer Aided Chemistry (CAChe v. 4.9, Fujitsu Inc.), running on a Blue and White G3/G4 Apple Macintosh computer. The partition coefficient for the octanol/water system was calculated with ClogP package of CAChe.

Results and Discussion

Anthracene derivatives are investigated for their interaction with DNA by calorimetric, spectroscopic, and computational studies. These show a strong role of hydroxyethyl side chains in the binding interaction. First, the binding data are presented, and these are followed by characterization studies.

ITC Studies. Titration of a solution of MEA (**2**) into a solution of CT DNA indicated exothermic binding (Figure 1A), and addition of the ligand solution to DNA resulted in prompt heat release. Binding was saturated when the ratio of the concentration of the probe to that of DNA exceeded 0.5. Titration of the ligand into buffer (no DNA) indicated small, exothermic heat of dilution (top curve, Figure 1A). Dilution of CT DNA with buffer resulted in negligible heat release (data not shown).

The area under each peak of the ITC titration data was integrated, and the heats of dilution were subtracted. The net heat released during the titration was plotted as a function of [MEA]/[CT DNA] (Figure 1B), and the data have been fitted to a single-binding-site model (solid line, eq 1). Calculations using different initial values converged to the same set of binding parameters, consistently, and the corresponding binding parameters are collected in Table 1.

The enthalpy of binding of MEA, estimated from the above data, is -3.4 ± 1.5 kcal/mol, and exothermic binding of similar anthracene derivatives to CT DNA has been reported.[25] The best fit to the data indicated a binding constant of $(3.7 \pm 1.5) \times 10^5$ M^{-1}, and this value is an order of magnitude greater than the corresponding binding constants of AMAC, APAC, or N-Et-AMAC with CT DNA.[14] Note that these were measured under the same conditions of buffer and ionic strength. MEA differs from N-Et-AMAC in terms of the single OH group in the side chain, and hence the enhanced affinity of MEA is most likely due to the additional OH group in this probe. Note that the pK_a values of these two probes are expected to be similar (pK_a of MEA = 6.6 ± 0.3), and both ligands are expected to

TABLE 1: Thermodynamic Parameters Obtained for the Binding of Anthryl Probes to CT DNA from ITC Measurements (Tris Buffer, 298 K)

ligand	*n*	K_b (M^{-1})	ΔH (kcal/mol)	ΔS (cal/mol·K)	ΔG (kcal/mol)
N-Et-AMAC			0		
MEA	0.22 ± 0.08	$(3.7 \pm 1.5) \times 10^5$	-3.4 ± 1.5	1.5 ± 17	-7.2 ± 0.3
BMEA	0.15 ± 0.01	$(1.7 \pm 0.4) \times 10^5$	-5.5 ± 1.9	5.3 ± 6.3	-7.1 ± 0.2

be protonated to a similar extent, under the experimental conditions.

The above analysis also indicated that the binding site size ($1/n$) for MEA is ~4 base pairs, and this value is greater than what is anticipated from the neighbor exclusion principle.[14] Analysis of the data with smaller binding site sizes did not give good fits and hence is not valid.

Similar to the exothermic binding of MEA, titration of BMEA (169 μM) into CT DNA (100 μM) also indicated strong heat release (Figure 2A). Binding saturated at a ratio of probe to DNA concentrations of >0.35. Subtraction of the corresponding heats of dilution and subsequent analysis of the data indicated that the enthalpy of binding of BMEA is -5.5 ± 1.9 kcal/mol. This value is greater than that of MEA, or other anthracene probes reported earlier.[15] Analysis of the BMEA binding isotherm, as in the case of MEA, by single-binding-site model resulted in satisfactory fits to the data (Figure 2B), and these indicated a binding constant of $(1.7 \pm 0.4) \times 10^5$ M^{-1}. Note that ITC analysis indicates that the affinity of BMEA is smaller than that of MEA, but these values are nearly within experimental error.

Absorption Studies. The absorption spectra of MEA and BMEA (300−420 nm) undergo significant changes on binding to CT DNA, and these provided a convenient handle to characterize their interactions (Figure 3A,B). The absorption spectra of the probes were recorded in the presence of increasing concentrations of DNA, and the spectra have been corrected for small changes in volume during the titration. The extinction coefficients of the bound chromophore were determined from the *y*-intercepts of the half reciprocal plots of absorbance versus 1/[DNA] (Table 2).

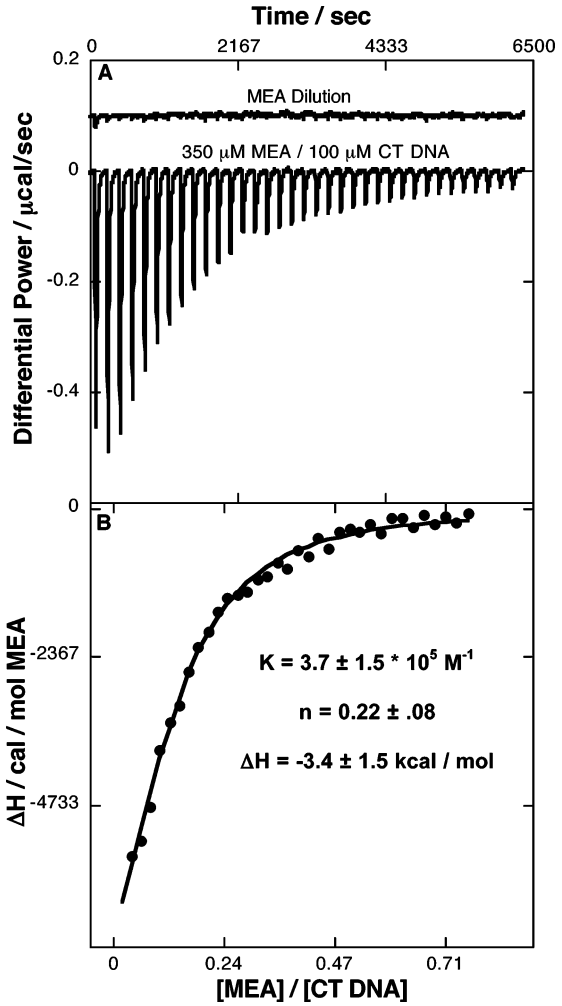


Figure 1. Isothermal titration calorimetry data for the binding of MEA (350 μM) to CT DNA (100 μM, 5 mM Tris, 50 mM NaCl, pH 7.2, 25 °C). (A) Heat released during the titration of CT DNA (100 μM) by successive additions of a concentrated solution of MEA (350 μM). The heat released during the dilution of MEA (no DNA) is shown separately. (B) The MEA/DNA binding isotherm from Figure 1A data; the solid line represents a fit according to the single binding site model.

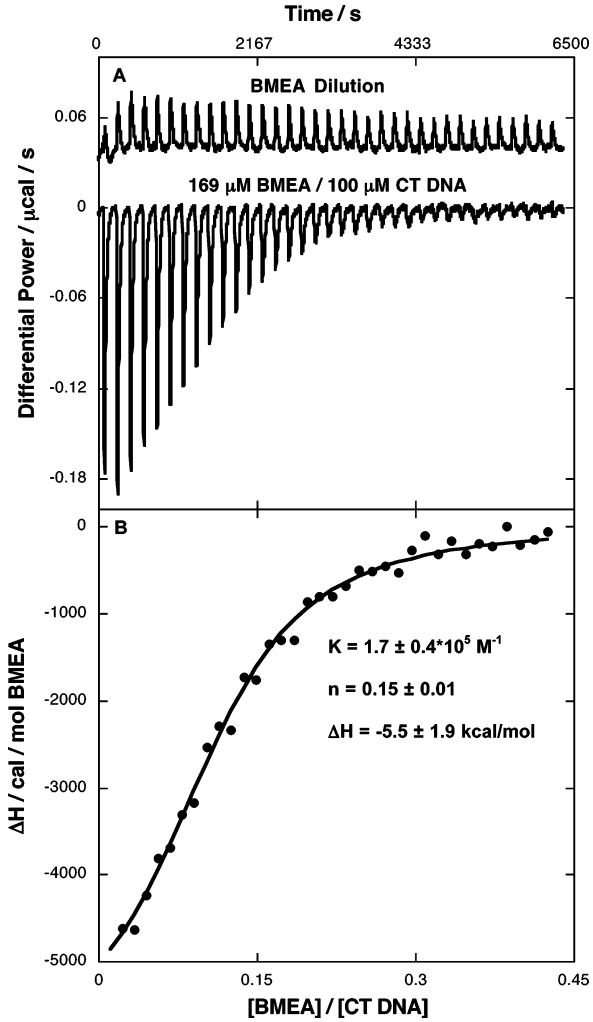


Figure 2. The isothermal titration calorimetric data for the binding of BMEA to CT DNA. (A) The heat released (top) when a concentrated solution of BMEA (169 μM) was titrated into a solution of CT DNA (100 μM, 5 mM Tris, 50 mM NaCl, pH 7.2, 25 °C). (B) The BMEA/CT DNA binding isotherm, where the solid line represents the best fit to the data, according to the single binding site model.

TABLE 2: Spectroscopic Data Obtained for the Probe/CT DNA Mixtures (Tris Buffer, 298 K)

probe	isosbestic points (nm)	λ_{max} free (nm)	λ_{max} bound (nm)	ΔE (cm^{-1})	%hypochromism (@ nm)	ϵ_{free} (M^{-1} cm^{-1}) (@ nm)	ϵ_{bound} (M^{-1} cm^{-1}) (@ nm)
MEA	394	387	396	650	67 (387)	6100 (367)	1700 (367)
BMEA	400, 384	394	404	630	76 (394)	6940 (373)	1800 (373)

The MEA/CT DNA absorption titrations, for example, indicated extensive broadening in the 300−420 nm region and a significant red shift (650 cm^{-1}). Absorbance at 394 nm was independent of the DNA concentration (isosbestic point), and the presence of the isosbestic point clearly indicates the smooth conversion of one chromophore to another (free and bound) as the titration progressed. The area under the absorption curve was drastically reduced (hypochromism) in the presence of excess of DNA, and the extent of hypochromism at 387 nm, for example, was 67% (Table 2). Such large hypochromism values were also noted when anthracene derivatives are intercalated in the helix.[14]

Similar to MEA, the absorption spectra of BMEA (12 μM) were recorded in the presence of increasing concentrations of CT DNA (0−132 μM) (Figure 3B), and these also indicated red-shifted (630 cm^{-1}) maxima as well as large hypochromism (76% @ 394 nm). Clear isosbestic points at 400 and 384 nm are also evident from the spectra. Note that the vibronic bands of BMEA/CT DNA are broadened extensively but they are clearly resolved, unlike in the case of MEA/CT DNA (Figure 3A).

The large hypochromism values and large red shifts are attributed to a strong electronic interaction between the π-electrons of the probe and those of the DNA bases.[26] The magnitude of hypochromism is expected to depend inversely on the distance of separation between the ligand and the nucleobases as well as the relative orientations of the interacting moieties.[27] The strengths of these electronic interactions of MEA and BMEA with DNA are comparable to those of APAC and N-Et-AMAC. The absorption characteristics of the anthracene probes, such as the above, have been assigned to the intercalative binding of anthracene probes (Chart 1).[15] On the basis of these reports, we suspect that MEA and BMEA also intercalate into the helix.

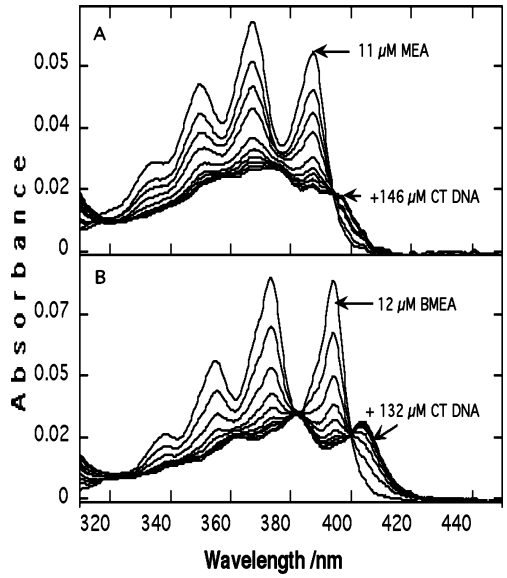


Figure 3. Absorption spectra of MEA and BMEA in the presence of increasing concentrations of CT DNA (5 mM Tris, 50 mM NaCl, pH 7.2, 25 °C): (A) MEA (11 μM) in the presence of CT DNA (0−146 μM), and (B) BMEA (12 μM) in the presence of CT DNA (0−132 μM).

TABLE 3: Intrinsic Binding Constants Estimated from the Absorption Titrations (Tris Buffer, 298 K)

probe	K_b (10^5 M^{-1})[a]	$\Delta G°$ (kJ/mol)[a]	K_b (10^5 M^{-1})[b]	$\Delta G°$ (kJ/mol)[b]
N-ET-AMA C	0.8 ± 0.05	-24 ± 1.0	0.12 ± 0.001	-23 ± 0.1
MEA	5.5 ± 0.1	-33 ± 0.1	1.0 ± 0.01	-28 ± 0.1
BMEA	8.0 ± 1.3	-34 ± 0.5	2.0 ± 0.01	-30 ± 0.1

[a] From Scatchard plot. [b] From von Hippel plot.

Intrinsic Binding Constants. The absorption titration data were analyzed to estimate the corresponding binding constants by using the Scatchard equation (eq 2). Binding isotherms were constructed from these data (Figure 4), and best linear fits to the data resulted in binding constants of 5.5×10^5 and 8.0×10^5 M^{-1} for MEA and BMEA, respectively. As an alternative to Scatchard analysis, the data were also analyzed using the McGhee and von Hippel eq 3.[21] All of the corresponding binding parameters are collected in Table 3.

Note that Scatchard analysis gave higher binding constants than von Hippel plots, but both indicated the same trends. According to both models, the binding constant of BMEA is larger than that of MEA, and at first sight this may appear to be due to higher charge on BMEA than MEA. The pK_a of $MEAH^+$ is 6.6 ± 0.3, and this value is expected to be comparable to the corresponding second pK_a of $BMEA(2H)^{2+}$. The first pK_a of $BMEA(2H)^{2+}$ is 5.9 ± 0.3, and the second amino group is unlikely to be protonated at pH 7.2.[28] Therefore, BMEA would not exist as a dicationic species at pH 7.2, and a significant fraction of both probes will be singly protonated. Therefore, the greater binding affinity of BMEA over MEA or that of MEA over N-Et-AMAC cannot be explained by their pK_a values. Note also that the binding affinities obtained from ITC are in reasonable agreement with

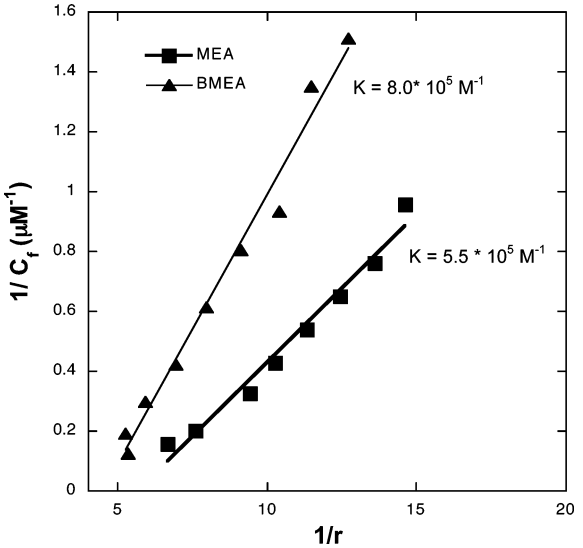


Figure 4. The binding isotherms for the anthracene probes (5 μM) with CT DNA (0−75 μM, 5 mM Tris, 50 mM NaCl, pH 7.2, 25 °C).

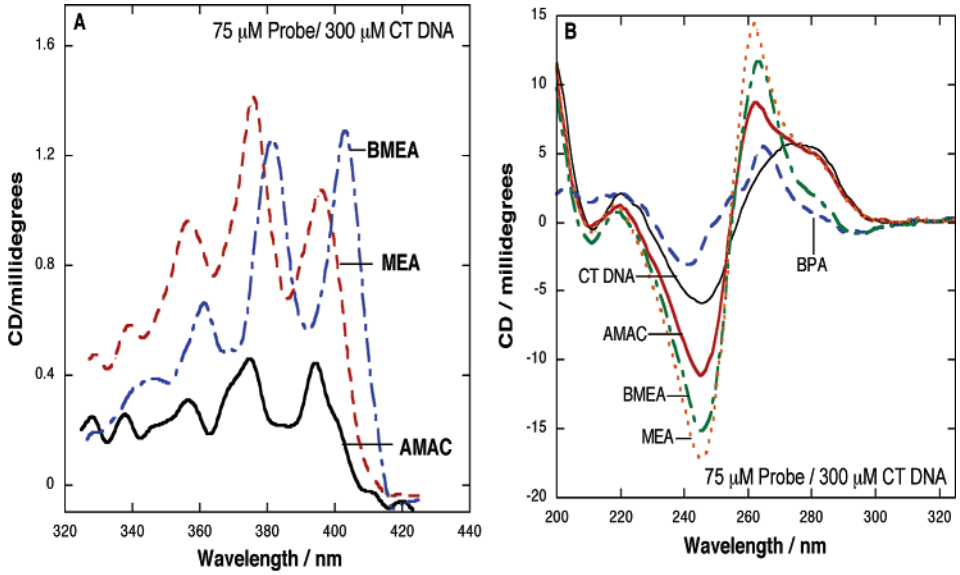


Figure 5. Induced circular dichroism spectra of the anthracene probes (75 μM) recorded in the presence of CT DNA (300 μM) (5 mM Tris, 50 mM NaCl, pH 7.2) in the near-UV region (A) and the UV region (B).

TABLE 4: Peak Positions of the ICD Spectra (Tris Buffer, 298 K)

probe	ICD peaks (nm)	UV−vis peaks (bound, nm)
MEA	398, 377, 358	very broad
BMEA	403, 381, 360	404, 382, 361

those of the absorption measurements but the latter are more accurate. The binding affinity data clearly illustrate the role of the side chains in controlling the overall affinities of the anthracene derivatives.

Induced Circular Dichroism (ICD) Studies. Binding of anthracene probes to DNA results in induced circular dichroism, and the ICD spectra indicate the orientation of the chromophore relative to the base pair dyad axis.[29] Positive ICD peaks are expected in the 320−400 nm region when the anthryl short axis is oriented perpendicular to the helix axis.[30]

Both MEA and BMEA showed positive ICD signals upon binding to DNA (Figure 5A), and these are consistent with the positive ICD spectra observed with other anthracene intercalators, including **1**, **4**, **5**, and **6**.[14] Note that the ICD peak positions of MEA and BMEA match with the absorption peak positions of the bound probes (Table 4), and these indicate that the ICD peaks correspond to the same chromophore that is responsible for the red-shifted absorption spectra in Figure 3.

In addition to the anthryl ICD data, the DNA CD at 280 nm is also an indicator of the probe binding modes. Unwinding of the DNA helix by classical intercalators, for example, results in strong decreases in the intensity of the 280 nm band.[31] Note that anthryl probes do not have a significant absorption in this region, and the contributions of any ICD signals in this region are negligible. The CD spectra in the 260−300 nm region are essentially due to the nucleobases. The intensity of the DNA CD at 280 nm is substantially reduced by BMEA (Figure 5B), and this observation is a strong indicator of classical intercalation. Similar to BMEA, a large reduction in 280 nm CD was also noted with another anthracene probe, BPA (R_1═R_2═CH_2−piperidine), and this reduction has been assigned to the classical intercalation of BPA.[15] In contrast to the behavior of these two probes, MEA did not cause such decreases in the 280 nm CD, and MEA may bind via partial intercalation. This is because the absorption spectral changes noted for MEA (Figure 3A) are not consistent with groove binding. Groove binding of anthracene probes did not produce any red-shifts in the absorption spectra.[15] Therefore, the two anthryl probes show specific differences in terms of their interaction with CT DNA.

Linear Dichroism Studies (LD). The above assignments of the binding modes are confirmed in LD studies. LD spectroscopy is an excellent tool to examine the orientations of chromophores bound to the DNA helix. The absorption transitions of the anthryl probes in the 320−400 nm region are short-axis polarized, while the 252 nm transition is long-axis polarized.[32] Classical intercalation of the probe results in both the long and the short axes of the probe oriented perpendicular to the helix axis, and this should give rise to negative LD signals at both absorption regions (∼252 and 320−400 nm). This was indeed observed at low loading of MEA and BMEA. The ligand LD at 252 nm is overlaid on the negative signal due to the DNA bases.

At low loadings, the LD spectra of the probes show negative peaks in the 252 and 320−400 nm regions (Figure 6). At high loadings, a small dip at 257 nm and negative LD in 320−400 nm region are noted. The spectra suggest intercalation of these probes. Intercalators usually stiffen the DNA, and that should lead to a larger DNA LD signal at 280 nm (anthrylene probes do not absorb in this region). LD data of a range of ligands suggest that major groove bound molecules tend to bend DNA, thus reducing the DNA LD, whereas the minor groove bound ligands stiffen the DNA.[33]

A comparison of the degree of change of the LD at the 257 (or 280) nm region indicates the differences in the binding modes of the probes. MEA significantly stiffens the DNA and has a small dip at 257 nm, appearing at DNA base:ligand ratios < ∼2.5:1. This small positive dip at 257 nm is assigned to a nonintercalative binding mode in which the anthracene long axis is less than 54.7° from the helix axis, perhaps stacking on the outside of the helix, or lying along a groove. It is important to note that this feature disappears when excess of DNA is present in the solution. BMEA also stiffens the DNA in a similar manner. The dominant binding mode for MEA can be concluded to be intercalation with its tail lying down the minor groove. The same is perhaps true for BMEA, although it has one side chain in each groove. Thus, the LD data lend strong support to the above conclusions.

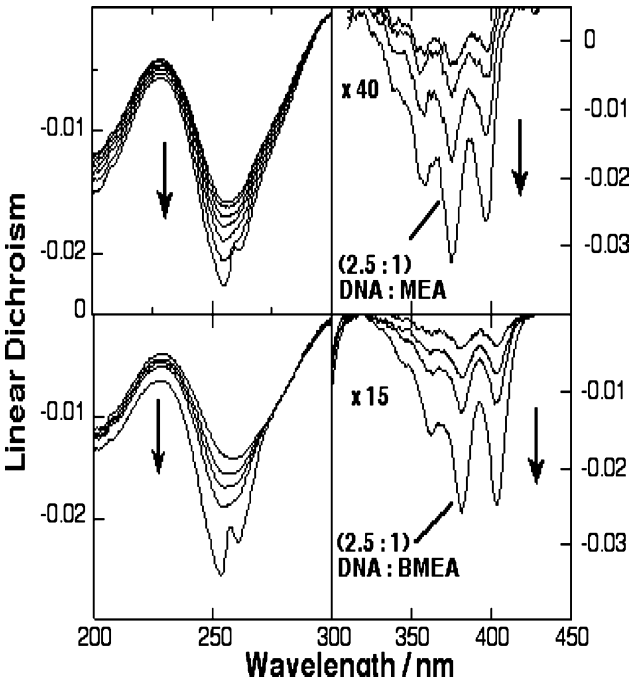


Figure 6. Linear dichroism spectra of CT DNA recorded in the presence of increasing concentrations of MEA and BMEA (1 mM sodium cacodylate, 50 mM NaCl, pH 6.9). In each set, the top spectra were obtained with CT DNA (160 μM CT DNA, no probe). Increments of the probe solutions were continually added to give decreasing ratios of DNA:probe concentrations (from 80:1 to 2.5:1) while keeping the DNA concentration constant by the addition of appropriate amounts of DNA solutions.

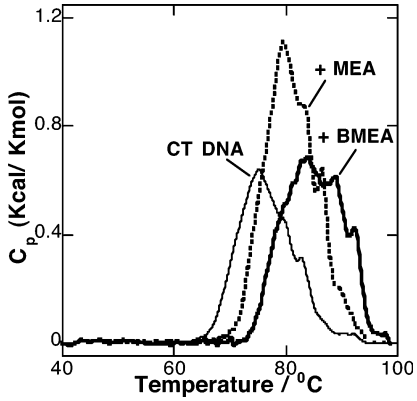


Figure 7. The differential scanning calorimetric thermograms of CT DNA (60 μM, thin line) and mixtures of the ligand/DNA complexes (5 mM Tris, 25 mM NaCl, pH 7.2). Significant improvements in the helix denaturation temperatures due to MEA (33 μM, dashed line) and BMEA (36 μM, thick line) are clear from the thermograms. The corresponding thermodynamic parameters are collected in Table 5.

DSC Studies. The thermodynamic characteristics of the helix denaturation were quantified by DSC studies. The DSC thermogram of CT DNA (60 μM) is shown in Figure 7 (thin line), which matches with the reported thermogram.[34] The helix denaturation is endothermic, and the DSC profile indicates one major transition around 75 °C as well as several minor transitions or satellite peaks. The satellite peaks are attributed to highly repetitive short base pair stretches of GC sequences, which denature at higher temperatures than other stretches of DNA.[34−36] Multiple DSC transitions of CT DNA are related to the heterogeneity of its base composition along the helix.[37]

The enthalpy change (ΔH) associated with the thermal denaturation is the integrated area under the DSC curve, and ΔH is model independent. On the other hand, ΔS was calculated assuming that the transition is reversible, and this assumption is valid/accepted in the literature.[36] Because of the heterogeneous nature of the base sequence in CT DNA, the renaturation of the single-stranded CT DNA is kinetically slow and not reversible on the DSC time scales. Even so, it is common to treat this transition as reversible and extract the corresponding thermodynamic parameters. The values thus obtained for CT DNA, from our data, are $T_m = 75.2$ °C, $\Delta H = 7.6 \pm 0.5$ kcal/mol, and $\Delta S = 22 \pm 5$ cal/mol·K (Table 5). These are in good agreement with the literature.[36]

TABLE 5: Thermodynamic Parameters Obtained for the Thermal Denaturation of CT DNA (60 μM), in the Absence and in the Presence of MEA (33 μM) or BMEA (36 μM) (Tris Buffer)

ligand	T_m (°C)	ΔH (kcal/mol)	ΔS (cal/mol·K)
CT DNA	75.2 ± 0.2	7.6 ± 0.5	22 ± 5
MEA	79.1 ± 0.1	11.2 ± 0.3	40 ± 5
BMEA	83.5 ± 0.7	10.0 ± 0.7	28 ± 2

Next, we examined the thermal denaturation profile of a mixture of MEA (33 μM) and CT DNA (60 μM) to further characterize the binding interactions. The data indicated an endothermic transition similar to that of CT DNA, but the profile is shifted to higher temperatures. Note that the satellite peaks are also shifted to higher temperatures in a corresponding manner. The enthalpy of denaturation of MEA/CT DNA estimated from these data (11.2 ± 0.3 kcal/mol) is significantly greater than that of CT DNA (7.6 ± 0.5 kcal/mol, Table 5). The helix stability, therefore, is increased by 3.6 ± 0.5 kcal/mol by the binding of MEA. Note that the extent of helix stabilization noted here is substantive and beyond the experimental error.

The DSC studies of BMEA/DNA mixture also indicated substantive changes in the DSC thermograms and increments in T_m (Figure 7A, Table 5). The helix stability is enhanced by 2.4 ± 0.5 kcal/mol, which is comparable to that observed with MEA. Note that similar increases in the ΔH values are also noted with other ligands.[25] Such stabilization was only noted with anthracene intercalators, while groove binding of anthracene ligands did not stabilize the helix.[15,38] Even though both intercalators as well as groove binders are known to improve helix stability, in the case of anthracene probes only intercalators show substantial improvements in helix stability.[15]

In addition to the above DSC studies, we also examined the influence of buffer characteristics on the above thermodynamic parameters. For example, the temperature coefficient of the pK_a of Tris buffer is substantial, while that of citrate is near zero.[39] We examined, therefore, the influence of buffer cations by using three buffers, while maintaining the same ionic strength. The thermodynamic parameters estimated for the denaturation of CT DNA in specific buffers showed measurable differences (Table 6). Despite these differences, the presence of BMEA resulted in similar increments in the T_m, ΔH, and ΔS values of CT DNA (Table 6). Therefore, the temperature dependence of the pK_a values of the buffer did not influence the helix stabilization endowed by BMEA.

The influence of the length of DNA helix on these parameters is also investigated by replacing the highly polymerized CT DNA by sonicated CT DNA. To minimize any buffer effects, the DSC data were collected with citrate buffer. The sonicated CT DNA did not show any differences from that of the intact CT DNA (Supporting Information). The corresponding thermodynamic parameters are collected in Table 6. These show that sonication of CT DNA did not have a significant effect on

TABLE 6: Thermodynamic Parameters Obtained for the Thermal Denaturation of CT DNA in the Presence and in the Absence of BMEA in Three Different Buffers (Citrate Buffer, 5 mM sodium citrate, 0.44 mM NaCl, pH 7.2; Tris Buffer, 5 mM Tris HCl, 25 mM NaCl, pH 7.2; and Phosphate Buffer, 10 mM Dibasic Potassium Phosphate, 7 mM NaCl, pH 7.2)

ligand	buffer	T_m/°C	ΔH (kcal/mol)	ΔS (cal/mol·K)
CT DNA	phosphate	72.9 ± 0.3	8.1 ± 0.4	23 ± 2
CT DNA/BMEA	phosphate	81.3 ± 0.8	8.6 ± 0.5	25 ± 3
CT DNA	citrate	67.6 ± 0.4	8.7 ± 0.8	26 ± 4
CT DNA/BMEA	citrate	76.3 ± 1.6	7.8 ± 1.1	23 ± 6
CT DNA (sonicated)	citrate	67.6 ± 0.3	8.3 ± 0.7	22 ± 4
CT DNA (sonicated)/ BMEA	citrate	76.2 ± 1.1	7.9 ± 0.6	23 ± 5

the denaturation thermodynamics of CT DNA, in the presence or absence of BMEA.

Computer Modeling. The role of hydrophobic interactions in the binding of the above anthracene probes was evaluated in computational studies. Attempts were made to correlate the intrinsic binding constants of the anthracene derivatives with the octanol−water partition coefficient (Log *P*, where *P* = [solute in octanol]/[solute in water]). This coefficient is a good predictor of the hydrophobicity, and the partition coefficients are calculated by ClogP package of CaChe. Anthracene derivatives with larger number of hydroxyethyl functions have lower Log *P* values due to the additional OH groups present in the side chain. As expected, the increased number of hydroxyl functions decreased the Log *P* values.

A plot of Log *P* values against the binding constants of anthracene probes is shown in Figure 8. Note that the correlation is poor, and significant deviations in specific cases are clear. Only in the cases of AMAC and N-ET-AMAC did the binding affinity increase with Log *P* value, but, in all other cases, the binding affinities increased with decrease in Log *P*. The linear fit has a negative slope, and this strongly suggests that affinity decreases with increased hydrophobicity. The binding constant does not increase with hydrophobicity. Therefore, other interactions play a strong role in determining the overall binding affinities of these molecules.

Given the fact that these probes are closely related to each other, the binding should be related to common structural features such as hydrophobicity, the net charge on the probe, or the number of H-bonding/hydroxyethyl groups present on the ligand. Because hydrophobicity does not appear to control binding, and affinity increases in a substantial manner, there should be other major contributors to the binding.

A plot of binding constants against the number of hydroxyethyl groups indicated a strong linear correlation (Figure 9). The plot consisted of five data points, including N-Et-AMAC and AMAC, and the linear fit to the data is excellent. The positive slope indicates that the binding constant increases as a function of the number of hydroxyethyl groups, and the slope gives the contribution of each hydroxyethyl group to the binding. This works out to be $\Delta K_{binding}$ of 4×10^5 M^{-1} per hydroxyethyl group. This is the first demonstration of such a strong correlation between the DNA binding affinities and the number of hydroxyethyl functions present on the probe.

Note that N-Et-AMAC serves as an excellent control in the above plot because this probe is similar in structure to MEA but lacks the single hydroxyl group. The intrinsic binding constant of MEA (5.5×10^5 M^{-1}) is approximately 6 times larger than the binding constant of N-Et-AMAC (0.88×10^5 M^{-1}). Thus, addition of a single hydroxyl group has a profound

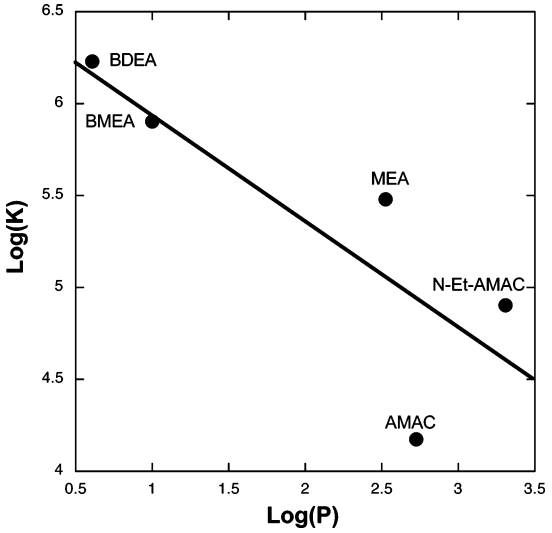


Figure 8. The plot of Log *K* versus Log *P* for the anthryl probes (the acronyms are defined in Chart 2). The solid line represents a linear fit to the data and indicated a poor correlation.

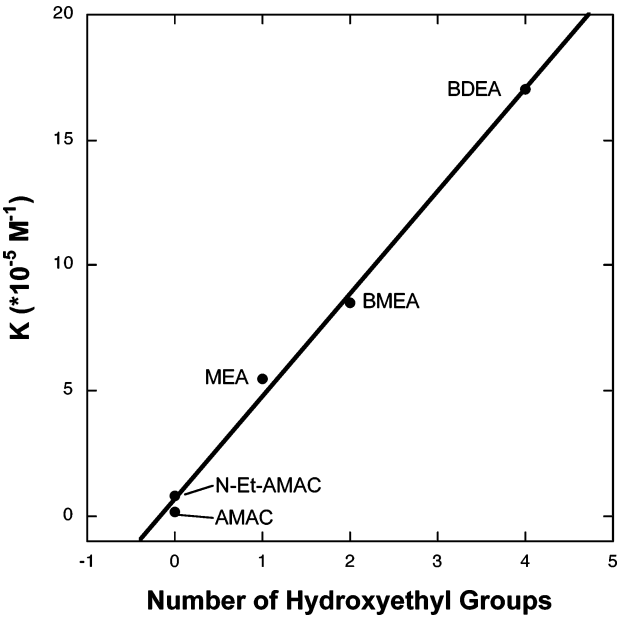


Figure 9. Plot of K_b as a function of the number of hydroxyethyl functions of the anthracene probes. The binding constants show a strong correlation with the number of hydroxyethyl groups present on the probe.

effect on the binding affinity. Even though the slope of the above plot includes contributions from the OH group as well as the CH_2–CH_2 unit, the contributions of CH_2–CH_2 to the binding affinity are either small or opposing to those of the OH group. A comparison of the affinities of AMAC and N-Et-AMAC (Figure 9) makes this point clear. The linear plot allows fair estimates of the DNA binding affinities of probes containing increasing number of hydroxyethyl functions and, perhaps, nonacidic OH groups.

Conclusions

Calorimetric, spectroscopic, and computational studies indicate that the anthracene probes bind to CT DNA with moderate to good affinities. The binding of both MEA and BMEA to CT DNA is exothermic (−3.4 to −5.5 kcal/mol), and binding is more exothermic with BMEA than with MEA. Because both probes are expected to be predominantly singly protonated at pH 7.2, this difference is likely due to the additional hydroxy-

ethyl group present in BMEA. Because the binding of AMAC or N-Et-AMAC to CT DNA is nearly thermoneutral, the enthalpy contribution due to each hydroxyethyl group can be deduced from these data.

The spectroscopic investigations indicate that the anthracene probes show bathochromic shifts ($\Delta E \approx 650$ cm^{-1}), extensive hypochromism (67−76%), strong positive CD peaks (300−450 nm region), and strong negative LD peaks in the 252 as well as the 300−450 nm region. These properties are also accompanied by helix stabilization ($\Delta T_m > 4-8$ °C) and significant increases in the enthalpy of helix denaturation ($\Delta\Delta H \approx 2.4-3.6$ kcal/mol).

These spectral and stability characteristics were identified with the intercalative binding of the anthracene probes but not groove binding.[15] Several anthracene probes bind via this mode, while there was also another binding mode that was associated with much weaker hypochromism (~30%), no shifts in the absorption spectral maxima, very weak ICD spectra in the 320−400 nm region, and no changes in the T_m of the ligand−DNA complex. Multiple binding modes are possible, but the spectral data indicate significant contributions of intercalation of both MEA and BMEA.

In case of BMEA, the DNA CD at 280 nm is reduced substantially. This reduction was attributed to the unwinding of the helix by the probe.[40] This interpretation was also supported by another example, BPA (R_1=R_2=CH_2−piperidine·2HCl), which also decreases the DNA CD at 280 nm to a significant extent. Therefore, we expect that intercalation of BMEA contributes to significant unwinding of the helix while intercalation of MEA is partial and does not unwind the helix substantially.

The linear plot in Figure 9 shows a strong correlation, and it will be useful to predict the binding affinities of similar molecules with hydroxyethyl substituents. This correlation also implies that the binding free energy varies nonlinearly among this set of molecules, and the underlying reasons for this behavior will be investigated in future studies. Meanwhile, the strong linear correlation and the slope of 4×10^5 M^{-1} per hydroxyethyl group are useful in the design of high-affinity DNA binding ligands.

Acknowledgment. C.V.K. thanks the National Science Foundation (DMR-0300631) for the financial support of this work.

Supporting Information Available: This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645−650. Helene, C.; Thuong, N. T.; Saison-Behmoaras, T.; Francois, J. C. *Trends Biotechnol.* **1989**, *7*, 310−15.

(2) Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 297−309.

(3) Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306−8311.

(4) Kimura, E.; Ikeda, T.; Shionoya, M. *Pure Appl. Chem.* **1997**, *69*, 2187. Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**, *7*, 355. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *382*, 559.

(5) (a) Tanious, F. A.; Jenkins, T. C.; Neidle, S.; Wilson, W. D. *Biochemistry* **1992**, *31*, 11632−11640. (b) Anthraquinone derivatives are well-known antitumor agents, see: Zee-Cheng, R. K. Y.; Cheng, C. C. *J. Med. Chem.* **1978**, *21*, 291. (c) Cheng, C. C.; Zee-Cheng, R. K. Y. In *Progress in Medicinal Chemistry*; Ellis, G. P., West, G. B., Eds.; Elsevier: Amsterdam, 1983.

(6) Carter, S. K.; Rall, D.; Schein, P.; Davis, R. D.; Wood, H. B.; Engle, R.; Davignon, J. P.; Venditti, J. M.; Schepartz, S. A.; Murray, B. R.; Zubrod, C. G. *Psuedourea*; National Cancer Chemotherapy Institute Clinical Brochure No. NSC56054, p 2.

(7) Pittillo, R. F.; Woolley, C. *Appl. Microbiol.* **1969**, *18*, 519−521.

(8) Becker, H. C.; Norden, B. *J. Am. Chem. Soc.* **1999**, *121*, 11947−11952.

(9) Wilson, W. D.; Tanious, F. A.; Watson, R. A.; Barton, H. J.; Strekowska, A.; Harden, D. B.; Strekowski, L. *Biochemistry* **1989**, *28*, 1984−1992. Wunz, T. P.; Dorr, R. T.; Alberts, D. S.; Tunget, C. L.; Einspahr, J.; Milton, S.; Remers, W. A. *J. Med. Chem.* **1987**, *30*, 1313−1321.

(10) Kumar, C. V.; Asuncion, E. H. *J. Am. Chem. Soc.* **1993**, *115*, 8547−8553.

(11) Kumar, C. V.; Asuncion, E. H. *Chem. Commun.* **1992**, 470−472. Rodger, A.; Taylor, S.; Adlam, G.; Blagbrough, I. S.; Haworth, I. S. *Bioorg. Med. Chem.* **1995**, *3*, 861.

(12) Kumar, C. V.; Tan, W. B.; Betts, P. W. *J. Inorg. Biochem.* **1997**, *68*, 177−181.

(13) Rodger, A. *Methods Enzymol.* **1993**, *226*, 232−258. Norden, B.; Kubista, M.; Kurucsev, T. *Q. Rev. Biophys.* **1992**, *25*, 51−171.

(14) Kumar, C. V.; Asuncion, E. H.; Tan, W. B. *Tetrahedron* **2000**, *56*, 7027−7040. Modukuru, N. K.; Snow, K. J.; Perrin, B. S., Jr.; Thota, J.; Kumar, C. V. *J. Phys. Chem. B* **2005**, *109*, 11810−11818. Modukuru, N. K.; Snow, K. J.; Perrin, B. S., Jr.; Bhambhani, A.; Duff, M.; Kumar, C. V. *J. Photochem. Photobiol., A* **2006**, *177*, 43−54.

(15) Tan, W. B.; Bhambhani, A.; Duff, M. R.; Rodger, A.; Kumar, C. V. *Photochem. Photobiol.* **2006**, *82*, 20−30.

(16) Robillard, M. S.; Galanski, M.; Zimmermann, W.; Keppler, B. K.; Reedijk, J. *J. Inorg. Biochem.* **2002**, *88*, 254−259. Zee-Cheng, R. K. Y.; Podrebarac, E. G.; Menon, C. S.; Cheng, C. C. *J. Med. Chem.* **1979**, *22*, 501−505.

(17) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor: New York, 1982; p 458.

(18) Wang, C.; Delcros, J. G.; Biggerstaff, J.; Phanstiel, O. *J. Med. Chem.* **2003**, *46*, 2663−2671.

(19) Wunz, T. P.; Dorr, R. T.; Alberts, D. S.; Tunget, C. L.; Einspahr, J.; Milton, S.; Remers, W. A. *J. Med. Chem.* **1987**, *30*, 1313−1321.

(20) Breslauer, K. J.; Freire, E.; Straume, M. *Methods Enzymol.* **1992**, *211*, 533. Sturtevant, J. M. *Annu. Rev. Biochem.* **1987**, *38*, 463−88.

(21) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469−486.

(22) Rodger, A. Linear Dichroism. In *Methods in Enzymology*; 1993; Vol. 226, pp 232−258.

(23) Refer for the partial specific molar volume of CT DNA: Durchschlag, H., Hinz, H. J., Eds. *Thermodynamic Data for Biochemistry and Biotechnology*; Springer-Verlag: New York, 1986.

(24) Doty, P.; McGill, B. B.; Rice, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 432.

(25) (a) Naga, K. M.; Snow, K. J.; Perrin, B. S.; Thota, J.; Kumar, C. V. *J. Phys. Chem. B* **2005**, *109*, 11810−11818. (b) Modukuru, N. K.; Snow, K. J.; Perrin, B. S., Jr.; Bhambhani, A.; Duff, M.; Kumar, C. V. *J. Photochem. Photobiol., A* **2006**, *177*, 43−54. (c) Tan, W. B.; Bhambhani, A.; Duff, M. R.; Rodger, A.; Kumar, C. V. *Photochem. Photobiol.* **2006**, *82*, 20−30.

(26) Long, E. C.; Barton, J. K. *Acc. Chem. Res.* **1990**, *23*, 271.

(27) Reinert, K.-E. *J. Mol. Biol.* **1972**, *72*, 593.

(28) Bisell, R. A.; Calle, E.; de Silva, A. P.; de Silva, S. A.; Gunaratne, H. Q. N.; Habib-Kiwan, J.-L.; Peiris, S. L. A.; Rupasinghe, R. A. D. D.; Samarasinghe, T. K. S. D.; Sandanayake, K. R. A. S.; Soumillion, J.-P. *J. Chem. Soc., Perkin Trans. 2* **1992**, *9*, 1559−1564.

(29) Ismail, M. A.; Sanders, K. J.; Fennell, G. C.; Latham, H. C.; Wormell, P.; Rodger, A. *Biopolymers* **1998**, *46*, 127−143.

(30) Ardhammar, M.; Norden, B.; Kurucsev, T. DNA−Drug Interactions. In *Circular Dichroism*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley−VCH: New York, 2000; pp 741−768. Lyng, R.; Hard, T.; Norden, B. *Biopolymers* **1987**, *26*, 1327.

(31) Lyng, R.; Hard, T.; Norden, B. *Biopolymers* **1987**, *26*, 1327−1345.

(32) Rodger, A.; Norden, B. *Circular Dichroism and Linear Dichroism*; Oxford University Press: Oxford, UK, 1997; pp 2−8.

(33) Rodger, A.; Blagbrough, I. S.; Adlam, G.; Carpenter, M. L. *Biopolymers* **1994**, *34*, 1583. Meistermann, I.; Parkinson, A.; Vidler, D. S.; Haworth, I. S. *Chirality* **2000**, *12*, 221. Rodger, A.; Parkinson, A.; Best, S. *Eur. J. Inorg. Chem.* **2001**, *9*, 2311.

(34) Klump, H. In *Biochemical Thermodynamics*; Jones, M. N., Ed.; Elsevier: Amsterdam, 1988; p 100.

(35) (a)Klump, H. *Ber. Bunsen-Ges. Phys. Chem.* **1987**, *91*, 2018. (b) Klump, H.; Herzog, K. *Ber. Bunsen-Ges. Phys. Chem.* **1984**, *88*, 20.

(36) Kurnit, D.; Shafit, B.; Maio, J. *J. Mol. Biol.* **1973**, *81*, 273.

(37) Petraccone; et al. *Thermochim. Acta* **2004**, *418*, 47−52.

(38) (a) Berman, M, H.; Young, P. R. *Ann. Rev. Biophys. Bioeng.* **1981**, *10*, 87−114. (b) Leng, F.; Chaires, J. B.; Waring, M. J. *Nucleic Acids Res.* **2003**, *31*, 6191−6197. (c) Leng, F.; Priebe, W.; Chaires, J. B. *Biochemistry* **1998**, *37*, 1743−1753.

(39) Perrin, D. D.; Dempsey, B. *Buffers for pH and Metal Ion Control*; Chapman and Hall Laboratory Manuals: London, 1974.

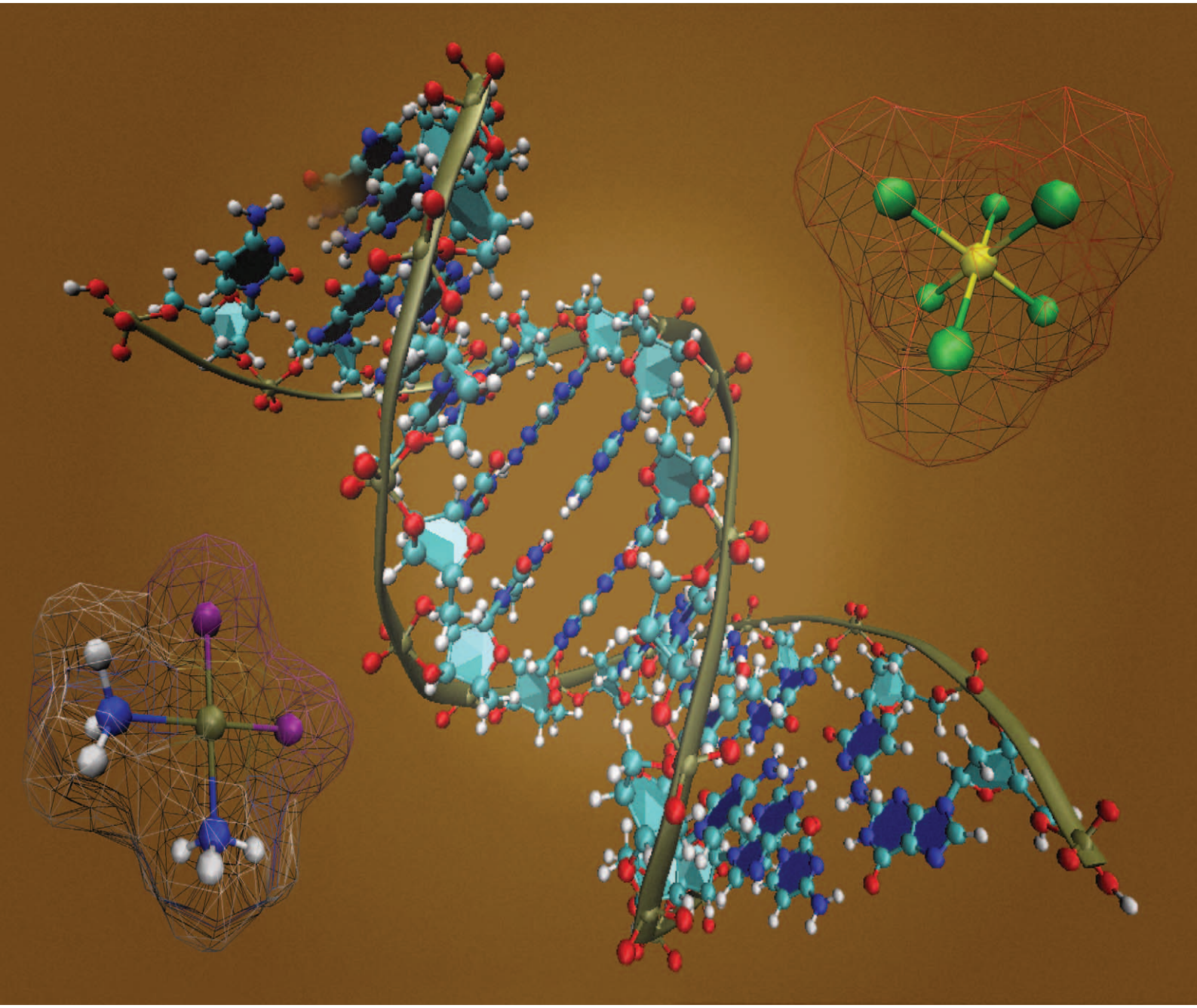
(40) Lyng, R.; Hard, T.; Norden, B. *Biopolymers* **1987**, *26*, 1327−1345.

Chem Soc Rev

Chemical Society Reviews

www.rsc.org/chemsocrev

Volume 36 | Number 3 | March 2007 | Pages 441 – 572



ISSN 0306-0012

RSC Publishing

TUTORIAL REVIEW

Adair D. Richards and Alison Rodger

Synthetic metallomolecules as agents for the control of DNA structure

Synthetic metallomolecules as agents for the control of DNA structure

Adair D. Richards[a] and Alison Rodger*[b]

Received 4th July 2006
First published as an Advance Article on the web 17th October 2006
DOI: 10.1039/b609495c

This *tutorial review* summarises B-DNA structure and metallomolecule binding modes and illustrates some DNA structures induced by molecules containing metallic cations. The effects of aquated metal ions, cobalt amines, ruthenium octahedral metal complexes, metallohelicates and platinum complexes such as *cis*-platin are discussed alongside the techniques of NMR, X-ray crystallography, gel electrophoresis, circular dichroism, linear dichroism and molecular dynamics. The review will be of interest to people interested in both DNA structure and roles of metallomolecules in biological systems.

1. Introduction to DNA

In 1868 Miescher[1] first isolated cell nuclei from pus cells and observed the presence of a phosphorus-containing compound. The resulting nucleic acids were found on hydrolysis to yield the purine bases: adenine and guanine and the pyrimidine bases: thymine and cytosine (Fig. 1). In 1950, Chargaff observed that the molar ratio of total purines to total pyrimidines was close to one to one, and that in deoxyribonucleic acid (DNA), the amounts of adenine and thymine were equivalent and of guanine and cytosine were also equivalent.[2] Three years later, Watson and Crick proposed a DNA structure made up of two helical chains, consisting of phosphate diester groups joining deoxyribofuranose residues, held together by purine and pyrimidine bases.[3] The combination of the base and sugar is called a nucleoside, which upon addition of a phosphate group generates a nucleotide.

Watson and Crick also suggested that the two helical chains were joined together in pairs, a single base from one being hydrogen bonded to a single base from the other chain and, using Chargaff's rule, that one base had to be a purine (A or G) and the other a pyrimidine (T or C).[3] Furthermore, the normal pairings of bases are adenine with thymine and guanine with cytosine. The pairs of bases, being planar, can be stacked one above the other. The molecule is therefore represented as a spiral staircase with the base pairs forming the steps. DNA exists in biological systems mainly in so-called B-form which is a right-handed helical structure (Fig. 2)[4] where the base pairs are perpendicular to the helix axis. Under idealised conditions, the diameter of the B-DNA helix is 2 nm, its pitch (the distance a helix rises along its axis per turn) is 3.4 nm and since there are 10 base pairs in each turn of the helix, there is a distance of 0.34 nm between each base pair. The pairing of the non-symmetric bases and their stacking, at approximately a 36° torsion angle (the angle between base pairs observed in the plane perpendicular to the helix axis),

[a]*Molecular Organisation and Assembly in Cells Doctoral Training Centre, University of Warwick, Coventry, UK CV4 7AL*
[b]*Department of Chemistry, University of Warwick, Coventry, UK CV4 7AL. E-mail: A.Rodger@warwick.ac.uk; Fax: 44 (0)2476 575795; Tel: 44 (0)2476 574696*

Adair Richards

Adair Richards was educated at the University of Warwick and received his BSc in Mathematics in 2003 and then attained a distinction in his MSc in Mathematical Biology and Biophysical Chemistry in 2004. He is now studying for his PhD, also at the University of Warwick, under the supervision of Albert Bolhuis and Alison Rodger in collaboration with Mike Hannon of the University of Birmingham. His project involves researching the interactions of supramolecular metallomolecules with biological systems at the sub-cellular level.

Alison Rodger

Alison Rodger graduated from Sydney University with the University Medal for Theoretical Chemistry in 1982. She completed her PhD in 1985 whereupon she took up a Beatrice Dale Research Fellowship at Newnham College, Cambridge together with an Overseas Scholarship from the Royal Commission for the Exhibition of 1851. She next spent six years in Oxford at St. Catherine's and St. Hilda's Colleges where she held a Glasstone Research Fellowship, moving to Warwick in 1994. In 2005 she became Professor of Biophysical Chemistry. Her laboratory focusses on developing and implementing UV/visible spectroscopic techniques (including circular and linear dichroism) to study the structure and function of biomacromolecules.

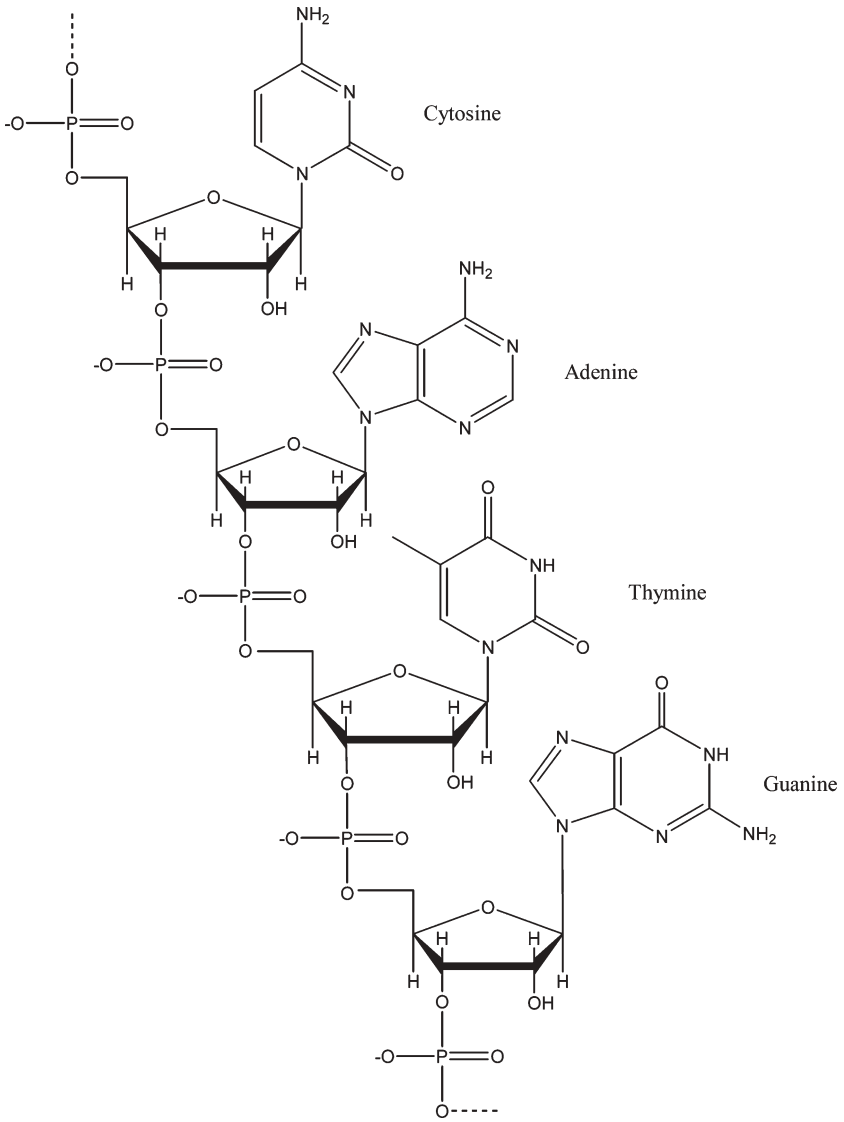


Fig. 1 The four DNA bases showing their connectivity to the phosphate-sugar backbone in DNA.

gives rise in B-DNA to a small (minor) groove and a large (major) groove[5] (Fig. 3). The realisation that the planar bases can associate in particular ways by means of hydrogen bonding was a major step in the understanding of the structure and function of DNA.[6]

It does not take much reading to recognise that DNA is a key molecule for the operation of almost any biological system, since it contains the instructions for making the correct proteins. However, it is not simply a genetic code to be read. One of the big challenges of today is to work out how transcription of certain pieces of DNA occurs at the correct times and to the correct extent to maintain a viable system. It is reasonable to conclude that DNA structures are important in this control. Thus, once one has established the 'norm' (usually B-DNA) for a system it becomes extremely attractive to find out about the deviations from this standard. In the case of DNA there are different polymorphs such as A-DNA and Z-DNA, each of which has roles to play. We also know there is a wide range of DNA structures without which biology would not be as we know it. Holliday junctions (mobile junctions

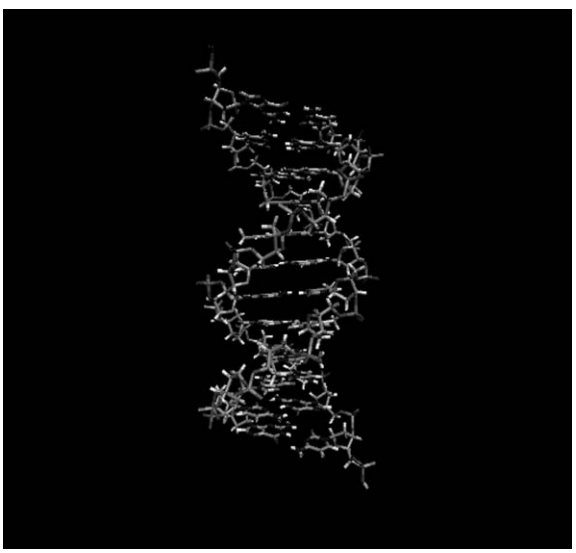


Fig. 2 Schematic representation and space-filling model of B-DNA (created using VMD[4]).

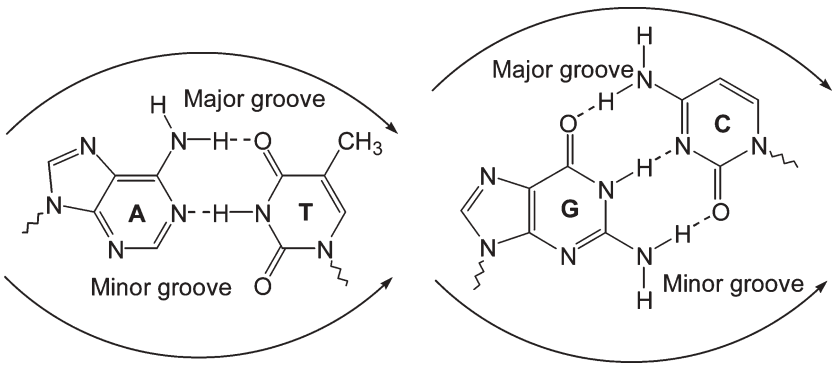


Fig. 3 Adenine–thymine and guanine–cytosine base pairs showing the minor and major grooves. The wiggly bonds indicate the links to the sugar backbone.

between four strands of DNA),[7] for example, look like a mess of spaghetti until one sees a colour-coded ribbon version at which point it becomes apparent that this is a very clever way of exchanging genetic material during cell meiosis.

Rather than grasping the whole topic of DNA structure and all its roles, in this review a brief overview of how some metallobiomolecules can be used to control DNA structure is given. Metal complexes are particularly attractive systems to study as they are usually cationic, which ensures their attraction to DNA. Furthermore, the fact that, in most such complexes mentioned the metal holds ligating organic molecules firmly in place means that the shape and surface features, *e.g.* hydrogen bond donors or acceptors, can be used to affect the DNA–metallocomplex interaction. In the text that follows some well-known examples are summarised which illustrate the variety of effects metallomolecules can have on DNA.

2. Modes of molecular binding to DNA

Despite the view often formed of DNA as a naked train track stretching out to infinity (on a molecular scale) this is not realistic. In practice DNA is wrapped up tightly like a ball of anionic (negatively charged) knitting wool. At the very least it needs some cationic species there to enable it to hold together. Much of this role in cell nuclei is provided by histone proteins, however, metal salts and other small molecules such as polyamines (see below) contribute significantly to this task. When one wishes to consider DNA structure control by specific molecules, it is important to consider where and how they bind to the DNA. Often the process of molecular recognition involves conformational adjustments on the part of the interacting species, and the DNA helix may end up considerably distorted. DNA distortion induced by molecules containing metals are the subject of this review. Three fundamentally different modes of DNA binding by metallomolecules can be identified: non-specific external association, (major and minor) groove binding, and intercalation between DNA bases.

2.1 External association

In 1968, Eichhorn and Shin discussed the ability of metal ions to interact with DNA and suggested two essential binding modes: external association with the negatively charged phosphate backbone and interactions with the electron donor groups of the bases.[8] (It should be noted that when we say 'metal ion', particularly in the context of transition metals, that the ion almost certainly has at least a solvation sphere and may also have explicitly coordinated (attached) solvent molecules making it a larger entity.) The factors influencing the backbone mode of binding are predominately electrostatic interactions and include the charge of the molecule, the ligand hydrophobicity and the total size of the ions. Metals can also bind covalently to the phosphate groups. One of the main consequences of external association of cations to the DNA backbone is that separate DNA duplexes associate together to form aggregates or condensed structures rather than repel each other.

2.2 Major and minor groove binding

In 1981, Gale *et al.*[9] noted that most small molecules that bind to B-DNA grooves do so either in or *via* the minor groove of the double helix, while DNA binding proteins or gene-targeted oligonucleotides interact with the major groove (Fig. 2 and 3). The reason for this is that many DNA binding molecules are small, fairly flat and cationic so the minor groove presents a better receptor for such molecules in terms of its size, flexibility, electrostatic potential and water bonding properties.[6] Goodsell and Dickerson later observed that minor-groove-binding drugs have a characteristic structure, usually containing two to five aromatic heterocycles linked by amide or vinyl groups with cationic groups at either ends.[10] Thus metal complexes do not fit neatly into the minor groove, though in some cases part of such a molecule may slot into the minor groove, *e.g.* ref. 11. The minor groove binder illustrated in Fig. 4 is stabilised on the DNA by van der Waals interactions with the walls and floor of the groove, as well as hydrogen bonds with the concave curvature of the inner surface of the molecule complementing the convex surface of the floor of the DNA minor groove. Such molecules have a preference for A–T

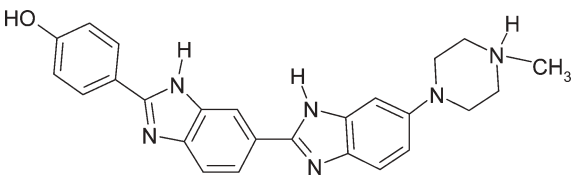


Fig. 4 Chemical structure of a typical minor groove binder Hoechst 33258.

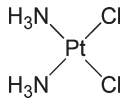


Fig. 5 Chemical structure of *cis*-platin.

rich sequences for which the groove has a deeper electrostatic potential and less steric hindrance. As a general rule, minor groove binding molecules have little effect on the structure of the DNA to which they bind, for example having little or no effect on DNA circular dichroism (though the spectrum may change if the ligand has a transition in the DNA region of the spectrum) or on the DNA orientation in a Couette flow linear dichroism cell (though the ligands themselves are being oriented as a result of binding).[12]

2.2.1 Platinum complexes. Despite the accepted preference of groove binding molecules for the minor groove, there are classes of small molecules which are accommodated by the major groove. For example, platinum complexes such as *cis*-platin (brand name Platinol, (Fig. 5), discovered by Rosenberg *et al.*[13]) and its second and third generation derivatives are major groove binders. They have proved to be the most extensively used anti-tumour agents invented to date. In 1970, the biological target of cisplatin was identified as DNA and it was shown to bind covalently to purine bases, in particular the major groove N7 of guanines.[14]

In the context of this article, a key feature of the favoured binding mode of *cis*-platin to DNA is that it binds to two neighbouring guanines on the same strand and induces unstacking of the bases at the Pt-binding site. This causes the DNA to kink by 53° (Fig. 6).[15] The first indication that *cis*-platin causes such a distortion was the crystal structure for the major *cis*-platin–DNA adduct, *cis*-[$Pt(NH_3)_2${d(pGpG)}], obtained in 1988[16] which showed the platinum metal centre co-ordinated in a square planar mode to two *cis*-ammine ligands and two guanine N7 atoms, in a manner that caused unstacking of the guanines.[16] The concomitant bending and local unwinding of the DNA double helix caused by platinum binding are important structural motifs; in recognition and differential processing of damaged DNA by repair enzymes or alternatively the binding of proteins which protect the lesion from repair — as in the case of *cis*-platin.[17]

Over the last few years, the field of platinum anti-cancer drug discovery has widened to polynuclear platinum complexes with Farrell *et al.* reporting di- and tri-nuclear platinum complexes (Fig. 7) showing some unique potential anticancer agent properties and activities against *cis*-platin-resistant tumours.[18]

2.2.2 Ruthenium complexes. Scientists are also looking at the possibility of using transition metals other than platinum as chemotherapy agents. Using different metals would offer possibilities such as having additional co-ordination sites, different oxidation states, and different ligand affinities, as well as providing the opportunity for electrochemical or photodynamic therapy (treatment with light-activated drugs *in vivo*).[19]

In particular, Alessio and co-workers have prepared a ruthenium(III) compound, imidazolium trans-imidazoledimethylsulfoxide tetrachlororuthenate, ImH[*trans*-$RuCl_4$(DMSO)Im], (Fig. 8) which is the first ruthenium-based compound to undergo clinical trials for its anti-metastasic properties.[20] It currently has successfully passed phase I clinical trials and entered phase II administered in combination with a primary antiblastic compound in the summer of 2006 in the Netherlands. Although the location of the molecule on DNA is not yet known, it does form inter-strand cross-links and appears to be a groove binder.

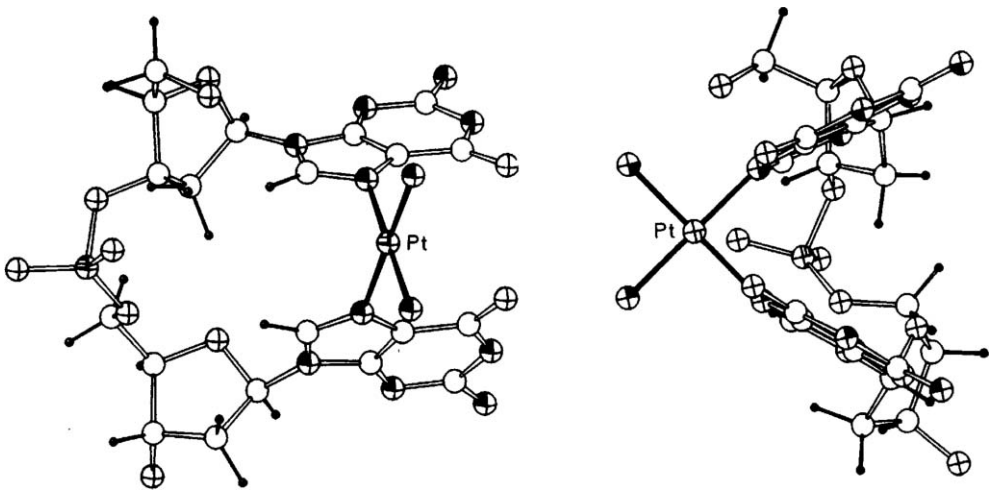


Fig. 6 DNA kinking induced by *cis*-platin binding.[15]

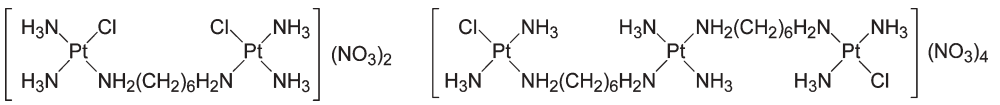


Fig. 7 [{*cis*-$PtCl(NH_3)_2$}$_2$$H_2N(CH_2)_6NH_2$]$Cl_2$ and BBR 3464.[18]

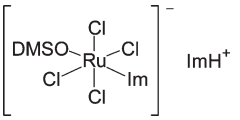


Fig. 8 ImH[*trans*-$RuCl_4$(DMSO)Im].[24]

Ruthenium(II) complexes have been widely used in binding studies for many years because of their stability (chemically, electrochemically, geometrically and enantiomerically) and also because of the sensitivity of their photophysical properties (luminescence enhancement and absorption hypochromism in the intense metal to ligand charge transfer bands) to binding interactions with DNA. An extensive debate can be found in the literature even over the binding modes of the simplest of the metal complexes with unsaturated ligands: ruthenium tris(1,10-phenanthroline), $[Ru(phen)_3]^{2+}$ (Fig. 9). One thing has been clear to all workers and that was that the two enantiomers of this complex interacted differently with DNA. In terms of structural effects, Hiort *et al.*[21] showed that the two enantiomers had different effects on the ability of DNA to be flow oriented. Λ did not alter the DNA orientation parameter until high loading whereas Δ reduced the orientation parameter by 50% at a DNA base : metal complex ratio of ~40 : 1. Coggan *et al.*[11] later concluded that at low loading, Δ bound with two of its chelates inserted into the minor groove, which distorts the DNA, resulting in it bending into a less effective flow orientation. Λ by way of contrast fits snugly and thus the major groove with one chelate partially inserted at the edge of the base pairs and does not bend or stiffen the DNA significantly.

2.2.3 Metal ions. Metal ions can have quite dramatic effects on DNA structure in addition to inducing aggregation or DNA bending (see above). For example, poly[d(G–C)]$_2$ DNA has been shown to change from the B form and stabilise in the Z form when the chloride salts of Mg^{2+}, Ca^{2+}, Cd^{2+}, Hg^{2+} or Zn^{2+} are added, even in relatively low concentrations.[22] Available X-ray data suggests that these ions bind in the major groove of B-DNA and cause the major groove to disappear, the minor groove to deepen and the helix to become left-handed.

2.2.4 Cobalt amine complexes. Another class of DNA structure altering metal complexes which have been extensively studied is cobalt amine complexes. The original motivation for this work stemmed from their apparent ability to mimic the behaviour of two DNA binding polycations: spermidine $[H_3N(CH_2)_3NH_2(CH_2)_4NH_3]^{3+}$ and spermine $[H_3N(CH_2)_3NH_2(CH_2)_4NH_2(CH_2)_3NH_3]^{4+}$ which are present in all cells.[23] The precise cellular functions of polyamines are unclear but they are known to charge neutralise over 40% of chromosomal DNA and they inhibit damage by reactive oxygen species, radiation, and alkylating agents. Further, polyamines stabilise DNA to heat and induce and stabilise A, B, Z (usually observed with alternating G–C sequences) and collapsed DNA structures such as toroidal condensates. They also readily induce structurally ill-defined DNA aggregates. Cobalt(III) amine transition metal complexes, *e.g.* $[Co(NH_3)_6]^{3+}$, (+)-$[Co(en)_3]^{3+}$, and (−)-$[Co(en)_3]^{3+}$ (en = ethylenediamine) mimic different aspects of the DNA interaction of the small polyamine molecules spermidine and spermine (references in ref. 24). From molecular modelling of $[Co(NH_3)_6]^{3+}$, (+)-$[Co(en)_3]^{3+}$, and (−)-$[Co(en)_3]^{3+}$ with DNA it was concluded that the cobalt amines do not bind in the minor groove of DNA, and their residence times in the major groove decreases in the order: spermine > $[Co(NH_3)_6]^{3+}$ > (+)-$[Co(en)_3]^{3+}$ > (−)-$[Co(en)_3]^{3+}$.[24] These amines bend DNA (as seen in gels and linear dichroism (LD)) as a precursor to condensing it and a key feature was concluded to be the presentation of a triangular NNN face to the DNA to optimise the bend-causing interaction. A similar geometric argument is required for the induction of the B→Z transition. Crystal structure data are consistent with the formation of five hydrogen bonds between $[Co(NH_3)_6]^{3+}$ and the major groove convex surface of Z-DNA, specifically to guanine O6 and N7 and phosphate oxygens, thus stabilising Z-DNA.[25] In a more extensive study the overall effectiveness of the amines in conferring thermal stability on the calf thymus (ct) DNA was found to proceed in the order:

Spermine > (+)-$[Co(en)_3]^{3+}$ ≈ (−)-$[Co(en)_3]^{3+}$ ≈ $[Co(NH_3)_6]^{3+}$ ≈ spermidine > $[Pt(NH_3)_4]^{2+}$ > $[Co(NH_3)_5H_2O]^{3+}$ ≈ $[Co(NH_3)_5Cl]^{2+}$ ≈ $[Co(NH_3)_5Br]^{2+}$ > *cis*-$[Co(en)_2NH_3Cl]^{2+}$ > $[Co(NH_3)_5(NO_2)]^{2+}$ ≈ *trans*-$[Co(NH_3)_4(H_2O)_2]^{3+}$

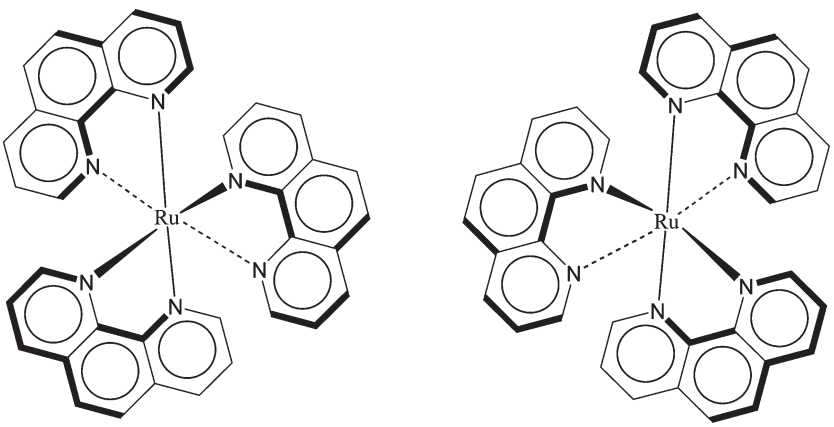


Fig. 9 The enantiomers of $[Ru(phen)_3]^{2+}$: Λ on left and Δ on right.

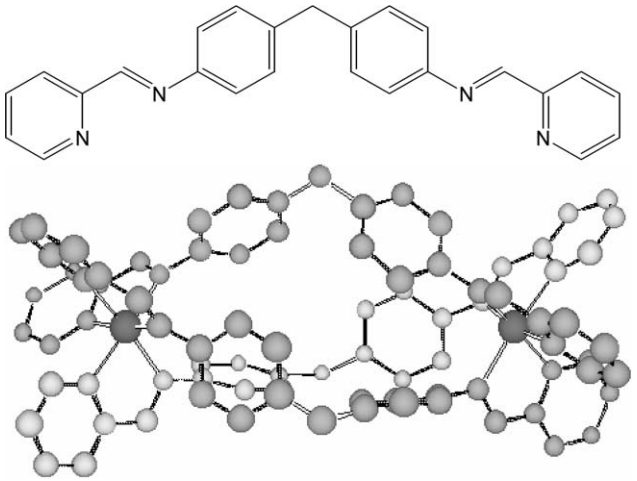


Fig. 10 Imine-based ligand used to bind metal co-factors such as Fe^{2+} to form a helicate.[27]

whereas the B→Z transition induction ordering is:

$$\text{Spermine} \approx [Co(NH_3)_6]^{3+} > (+)\text{-}[Co(en)_3]^{3+} \approx [Co(NH_3)_5(H_2O)]^{3+} > (-)\text{-}[Co(en)_3]^{3+} > [Co(NH_3)_5Br]^{2+} > [Co(NH_3)_5Cl]^{2+}$$

with the other amines failing to cause the transition under the 11 mM salt and temperature ramp conditions used. The DNA bending ranking is:

$$\text{Spermine} > [Co(NH_3)_6]^{3+} > (+)\text{-}[Co(en)]^{3+} > (-)\text{-}[Co(en)]^{3+} > [Co(NH_3)_5(H_2O)]^{3+} > \textit{trans}\text{-}[Co(NH_3)_4(H_2O)_2]^{3+} > \text{spermidine}$$

with the other amines having comparatively little effect. The most notable things about the above lists is that the rankings of the molecules varies with the structural effect being investigated. There is a general trend of reducing efficacy with reducing charge (the aquo complexes will be deprotonated at pH = 6.8) but the key conclusion from this work is that the molecular interaction surfaces are important for determining how metal complexes affect DNA structure. For example, the inability of the nitro and halo compounds to significantly bend the DNA (they are absent from the list above) leads to the conclusion that the bend-inducing binding mode requires deep penetration of the groove by hydrogen bonding donors. In accord with the hypothesis that a triangular face of amines facing the DNA is required to induce the B→Z transition, neither *trans*-$[Co(NH_3)_4(H_2O)_2]^{3+}$ nor $[Pt(NH_3)_4]^{2+}$ [26] are capable of inducing the B→Z transition (even in the absence of NaCl). The other compounds' ranking are consistent with their ability to face the DNA with NNN faces.[24]

2.2.5 Metallohelicates. A considerable amount of research in recent years has been focussed on the synthesis of supramolecular helicates with the purpose of having them interact with DNA to cause conformational changes. The term helicate in this context refers to helical coordination arrays containing at least one metal centre. The ligands that surround the metal ions must contain multiple coordination sites and flexibility in order to wrap around in a helical fashion and must include sufficient spacer groups to prevent adjacent binding domains binding to the same metal ion. The helicates may be 2, 3 or 4-stranded.

The simplest examples of our helicate are the tris-chelate complexes such as $[Ru(phen)_3]^{2+}$ mentioned above. More complicated examples of this type of compound involve the imine-based ligand systems[27] (Fig. 10) synthesized by Hannon and co-workers. Depending on the metal and the ligands, they form double or triple two-metal helicates with iron, copper, silver and nickel centres. The two metal binding sites of each individual chelate are prevented from binding to the same metal centre by the phenylenes in the ligand whilst the methylene group allows sufficient flexibility to wrap around in a helical manner.[27]

The most studied of the imine dinuclear iron-centred triple helicates has been shown to bind to DNA in the major groove and induce dramatic intra-molecular coiling.[28] Induced circular dichroism (CD) signals in the DNA and metal–ligand charge transfer region indicate that the helicate is binding in a single mode for helicate concentrations lower than 5 base pairs : 1 helicate, and positive LD signals for metal–ligand charge transfer transitions show that the helicate is binding in an oriented non-random fashion.[29,30]

Atomic force microscopy used in tapping mode showed that this helicate causes intramolecular supercoiling of free linearised DNA (Fig. 2.2.5.2) which is in distinct contrast to the commonly observed DNA intermolecular condensation and aggregation mentioned above with *e.g.* hexammine cobalt. The location of the helicate binding site was investigated using nuclear Overhauser effect spectroscopy (NOESY, see below for technique details).[31] There were a number of cross-peaks in the resulting spectra which indicated contact between the helicate and the major groove of DNA confirming that the helicate does indeed bind in the major groove. A later study[30] using CD and flow-LD spectroscopies noted the difference in

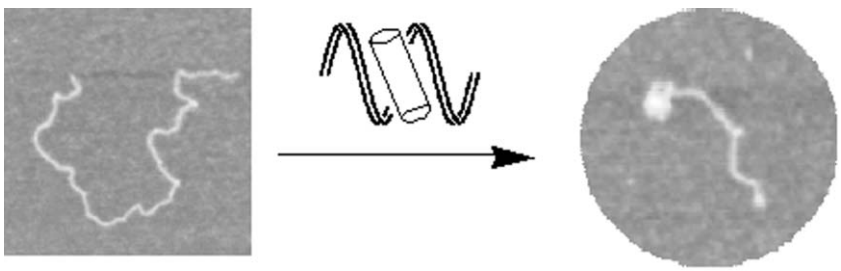


Fig. 11 AFM image of helicate inducing intramolecular DNA supercoiling.[28]

binding mode of each enantiomer. The conclusion was that the M enantiomer binds in the major groove whilst the P enantiomer may bind in the minor groove due to different complementarities of helicate-DNA surfaces.

Recently, this tetracationic helicate has been shown in an X-ray crystallography study to induce a junction with three-fold symmetry made up of three oligonucleotides with the helicate residing in the central cavity. The cavity induced in the DNA structure is a triangular shaped hydrophobic domain[28] formed by three double-strands of DNA that converge on one locality. This type of structure is unprecedented for a synthetic molecule but nicely illustrates how the symmetry possibilities afforded by metallomolecules can be used to template DNA structures. Whether such structures could be formed in the presence of polynomic DNA remains to be seen.

The basic message from this particular recent example is that we have in no way exhausted the range of structures that can be induced in DNA upon interaction with metallo-systems. The challenge is invariably how to recognise that a new structure has been formed and to identify its nature and function.

2.3 Intercalation between DNA base pairs

DNA intercalation was defined by Mainwaring *et al.* as the sandwiching of a molecule between two adjacent pairs of bases in the DNA double helix.[32] Intercalating ligands are characterised by the possession of an extended electron deficient planar aromatic ring system. Upon binding, they extend and unwind the deoxyribose–phosphate backbone and are stabilised by π–π stacking interactions with the planar aromatic bases.[33] Intercalation also leads to hydrodynamic changes in the DNA due to the decrease in twisting between the base pair layers, the lengthening of the DNA itself, the stiffening of the helix, and the decrease in mass per unit length. These effects are fully reversible upon removal of the intercalator as long as the DNA duplex structure is not destroyed by the process of removal. Lerman showed that a bound intercalator lies in a plane perpendicular to the helix axis, and that the perpendicularity of the base pairs to the helix is not significantly altered.[34]

Thus for a metal complex to be an intercalator it must either be planar or have an extended planar component which can slot between base pairs. In 1978 Lippard *et al.* published details of several planar platinum-based metallo-intercalators (Fig. 12) whilst seeking to understand the binding modes of the genre of platinum-based anti-tumour agents that were beginning to emerge (references within ref. 16). Another class of potential metallo-intercalator is the cationic porphyrins, which combines readily with metal centres in their central cavity and exhibits a strong binding affinity for DNA. There have been three binding methods described in the literature for such molecules (references within ref. 35): self-induced external auto-aggregation along the DNA helix axis, binding in either the major or minor groove of DNA and intercalation between base pairs. However, intercalation only occurs if the metal lies in the same plane as the porphyrin ring.

The ability of porphyrin compounds to bind to DNA has been used to bring other moieties into close proximity to DNA by using a porphyrin core as a molecular anchor. These moieties can be designed to increase binding affinity or to include chemical functionalities which can modify DNA structure *via* an independent method. Cu^{2+} compounds have been shown to operate as DNA cleavage agents[36] and recently dimetallo-copper-bipyridyl porphyrins[35] have been shown to bind to DNA by intercalation and external association and cleave DNA under certain experimental conditions. These molecules are formed by combining a copper–porphyrin ring which anchors the complex to DNA with a copper–bipyridinium moiety that hydrolyses phosphodiester bonds (Fig. 13). This rather dramatic structural effect on DNA is potentially useful either in the molecular biology laboratory or in drug design.

The less direct method of achieving intercalation of a metallo-system is illustrated by dipyridophenanzine-bi(phenanthroline) ruthenium(II), $[Ru(phen)_2dppz]^{2+}$ [37] (Fig. 14) where effectively a phenanthroline of $[Ru(phen)_3]^{2+}$ (Fig. 9) is extended by a planar aromatic group which prefers an intercalated binding mode. Its binding to DNA causes an elongation of the rod-like DNA molecule consistent with classical intercalation.[37]

3. Detecting DNA structure changes

The above discussion demonstrates that metal-containing molecules can have dramatic effects on DNA structure. However, how does one conclude that this is the case for a new molecule and determine what the effect is? There are various ways of measuring the deviation of the helix backbone from a straight line and various measures of DNA base displacement from the canonical B-DNA positions. Such measures really rely on having atomic level structural information from nuclear magnetic resonance spectroscopy (NMR), X-ray crystallography or molecular modelling studies and are

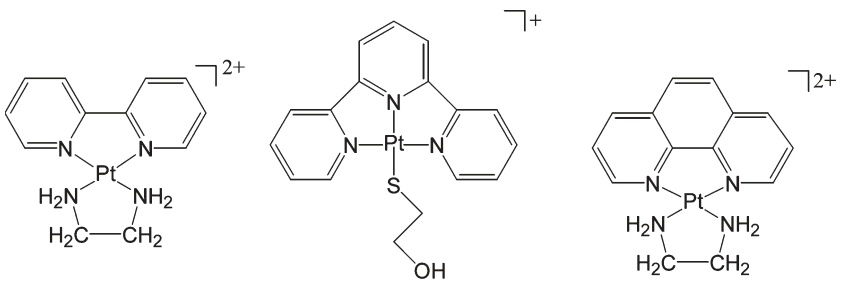


Fig. 12 Platinum complexes $[Pt(bipy)(en)]^{2+}$, $[Pt(terpy)(HET)]^{+}$ and $[Pt(o\text{-phen})(en)]^{2+}$.[12]

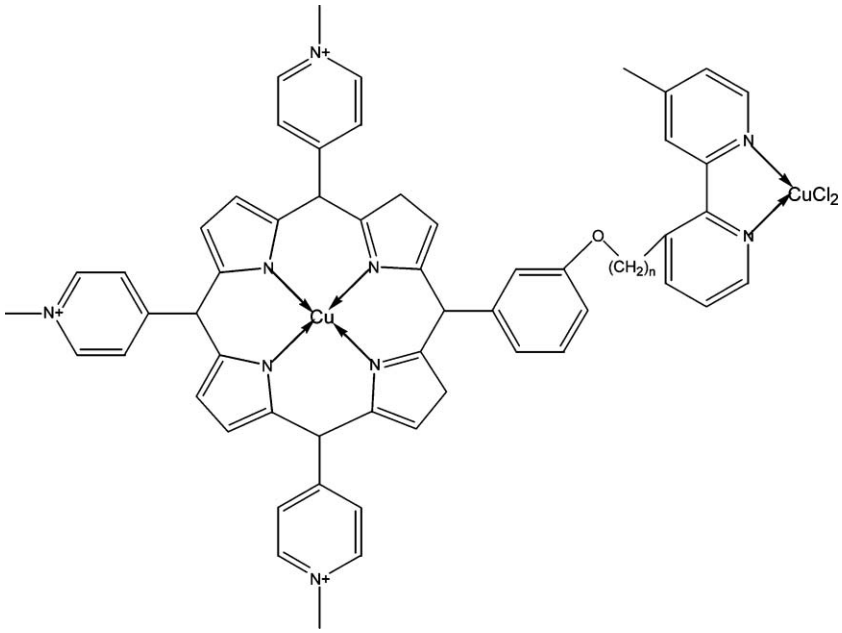


Fig. 13 3-[Cu-bpy-$(CH_2)_n$]–Cu-tMPyP.[35]

valuable when such data are available. Other techniques provide less detailed information but are usually easier to implement.

Given below is a description of the main techniques used to detect and analyse DNA structural changes in this context, alongside examples of how these techniques can be used. Note that the examples given are not intended to provide in-depth knowledge on all possible occasions where the technique may be adopted, but rather to give a variety of different applications where it has already been shown to be of use.

3.1 Nuclear magnetic resonance spectroscopy

NMR spectroscopy is a powerful technique that enables geometrical details of the structure of DNA on an atomic level in solution to be elucidated. In addition, changes in chemical shifts induced by temperature or concentration variation can provide information about DNA–complex association.

This technique relies on a property of nuclei known as spin, which is an intrinsic quantum physical property of a nucleus and has no direct analogue in the macroscopic world, but behaves like an angular momentum. When a nucleus is placed in an external magnetic field, its spin magnetic moment moves around the direction of the field at a frequency termed the Larmor frequency, which is directly proportional to the strength of the magnetic field experienced by the particle. Although nuclear spin has only a negligible effect on the chemical and physical properties of atoms, the energy levels or precession frequencies adopted by nuclei in different spin states are extremely sensitive to the individual chemical environment of a nucleus. Thus if one can measure the energy level gaps or precession frequencies one can in principle determine the local chemical environment of a nucleus of interest.

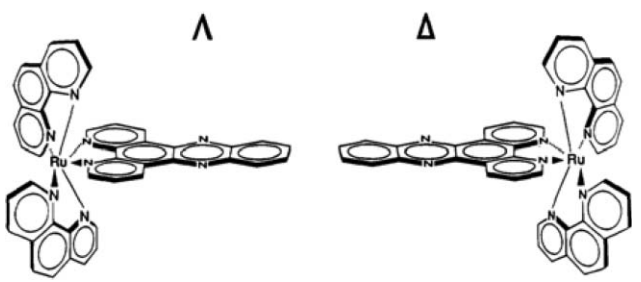


Fig. 14 The enantiomers of $[Ru(phen)_2dppz]^{2+}$.[37]

A radiofrequency pulse is applied to the system causing all the spin polarisations to rotate by 90°. When the pulse is switched off, a net spin polarisation exists perpendicular to the external field and since each individual spin precesses, the bulk magnetic moment does as well. The precession frequency of this bulk moment is the same as the Larmor frequency and it decays or 'relaxes' over time approximately exponentially and in an oscillating manner. An electric current is induced in a wire coil from this alternating magnetic field and is referred to as the NMR signal or free-induction decay.

NMR delivers information on a submolecular scale to a high resolution because of the inhomogeneous distribution of electrons in molecules which results in nuclei at different points within the molecule being subject to slightly different magnetic fields. Thus the Larmor frequency depends on the local chemical environment and leads to different nuclei having different chemical shifts. This is displayed as a resonance peak at a particular parts per million (10^{-6}) value on an (already Fourier-transformed) NMR spectrum. With metallo-systems binding to DNA, the extent of changes in chemical shifts for DNA protons upon ligand binding have often been used as an indication of where the ligand is located: more shift equals more perturbation hence indicating the binding site. However, one must take care as this simplistic approach is really predicated on there being little or no change to the DNA structure upon ligand binding. A better approach

is to use a range of alternative NMR techniques which give through-bond or through space-distances between atoms.

Nuclear Overhauser effect spectroscopy (NOESY) provides the ability to plot the through space connectivities between atoms with the strength of the signal observed depending on their distance. NOESY couplings arise because the use of a small radio frequency field at the Larmor frequency of one nuclear spin results in an enhancement of the magnetisation of some of the non-radiated nuclear spins.[31] Thus one can identify whether given atoms are near each other in three-dimensional structures. The NOESY data gives certain inter-atomic distances, which can be used as restraints in molecular modelling and one can therefore sometimes determine 3D structures.

NOESY data cannot be used to analyse the structure of a long DNA molecule as the peak-assignments on the resulting spectra becomes prohibitively complicated as the chemical shifts for various atoms overlap extensively. It can, however, be used to study oligonucleotides and molecules binding to oligonucleotides. An example of this is the use of molecular dynamics restraints obtained from NOESY data to determine features of the iron helicate mentioned above with a DNA decamer.[38] The accurate interpretation of NOESY spectra, or related techniques such as correlation spectroscopy (COSY)[31] and rotating frame Ovehauser enhancement spectroscopy (ROESY)[31] rely on molecular modelling to elucidate the correct molecular structure between multiple possible arrangements.

One application of NOESY was to show that the X-ray detected DNA conformational changes induced by *cis*-platin[15] were consistent with the behaviour observed when in solution. Assignment of the proton resonances were used to inform molecular dynamics simulations, from which was elucidated proton–proton and proton–phosphorus coupling constants. Analysis of these constants suggested a conformational change of one deoxyribose moiety leading to a kink in the DNA oligonucleotide. This is one of many examples of how the atomic-resolution information gathered by NMR can be used to provide valuable insights on the structural behaviour of DNA under the influence of metallomolecules.

3.2 X-Ray crystallography

X-Ray crystallography relies on Bragg's law which states that an X-ray beam reflecting from a surface layer of a crystalline material travels a shorter distance than those reflected by the inner layers. The beams are in-phase if the difference in these distances is an integer value of wavelengths of the incident radiation and hence produces an enhanced signal compared to when they are out-of-phase. Bragg suggested that the differences in distances relies upon the angle of incidence of the beam so that by changing this angle, a diffraction pattern can be built up which can then be Fourier-transformed and interpreted to give atomic-level structural information.[39] It is extremely attractive to have such data for metal complex–DNA systems, however, the challenge is that it is necessary to crystallise the sample. This is often not a trivial exercise, and always involves carefully chosen short DNA sequences. The biological significance of such structures is not clear.

It is important to note when studying and utilising the literature, particularly with X-ray and NMR data where the end result is a beautiful picture of the DNA–ligand system, that data determined for oligonucleotide systems are not necessarily applicable for pieces of DNA of the length of a gene or a genome. A simple illustration is provided by the three-way junction structure discussed above which is induced in oligonucleotides by a tetracationic dimetallohelicate molecule: it requires the appropriate base pairing to be available and the DNA strands to rearrange themselves from the standard form of B-DNA. There is also the fact that solid state crystalline structures derived from crystallisation buffers do not necessarily bear any relationship to those adopted in biological systems.

3.3 Gel electrophoresis

Persistence length is a useful concept when considering polymeric pieces of DNA. A working definition for the persistence length (Fig. 15) of a piece of DNA is taken to be the average length of DNA required for the helix backbone to bend 1 radian ($= 180/\pi = 57.3^\circ$). Typical B-DNA has a persistence length of about 150 base pairs. When the DNA is bent or made more flexible, this is reduced. Thus this parameter gives a measure of bending or flexibility of DNA. For example at the *cis*-platin binding site (Fig. 6), the persistence length is reduced to just a few base pairs. There is no really simple way to measure this fairly simple parameter. One method is to use AFM pictures of samples of DNA with and without the ligand of interest.[29] This is tedious and dependent on sampling methods to get a clear indication of what is happening. An alternative approach is to measure the changes in mobility of DNA in gels as a function of ligand loading on DNA.

Gel electrophoresis involves application of an electric field across the length of the gel causing the migration of charged molecules toward the oppositely charged electrode. At neutral pH, DNA is negatively charged so moves through the gel from cathode towards the anode. The electrophoretic mobility of a piece of DNA on a gel is dependent on its flexibility, size and charge.

If, when a molecule binds to DNA it affects the size or shape of the DNA, then it will affect its electrophoretic mobility. For example, an intercalator is expected to unwind the DNA,

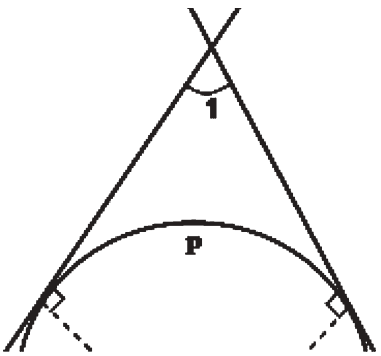


Fig. 15 Persistence length (p) can be thought of as the minimum length of DNA required such that the tangents to the helix axis form an angle of 1 radian.

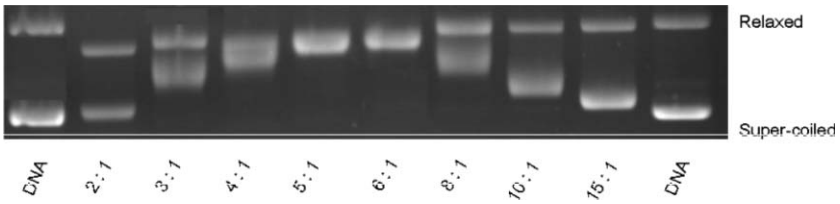


Fig. 16 Gel image of pBR322 DNA treated with *cis*-platin in DNA base : ligand ratios as indicated in the figure. The negatively supercoiled DNA sample has a small population of relaxed DNA. At 6 : 1 ratio both bands co-migrate; at higher ratios the DNA is being positively supercoiled.[42]

which lengthens and stiffens it. Using circular DNA of known supercoiling[40] the degree of unwinding per ligand bound can be determined. Whether the binding event is consistent with intercalation can then be determined. Ethidium bromide, the archetypical intercalator, unwinds DNA by an average of 26° per ligand (which is 10° less than the base–base twist in canonical B-DNA). The degree of unwinding induced by a given metal complex has been shown to indicate which binding mode it adopts:[41] mono-adduct platinum(II) complexes generally afford little unwinding, whereas bifunctional adducts such as *cis*-platin (unwinding angle 13°) have more effect, though significantly less than that of an intercalator (Fig. 16).[42]

An alternative application of gel electrophoreis has been described by Carle and Olson.[43] They used it as a method of fractionating DNA into size based fractions by applying a series of orthogonal electric fields on a gel. By altering the duration of the electric field pulses in each direction, a greater accuracy in the DNA separation for a given size range can be achieved. The technique can also be used to check for DNA charge reversal[44] which can occur when DNA is resolubilised in the presence of multivalent cations. Charge reversal is said to have occurred if the DNA migrates towards the cathode terminal.

3.4 Circular dichroism

Circular dichroism (CD), the difference in absorption of left and right circularly polarised light,[45] is uniquely sensitive to chirality or helicity of molecules. Thus DNA CD depends on the arrangement of the DNA bases and if this is changed, then the CD spectrum changes. The source of the chirality in DNA is the ribose sugar backbone of DNA, however, sugar CD signals are below 190 nm, so essentially inaccessible in laboratory-based CD machines. Between 190 and 300 nm, there are DNA signals observed from the coupling of the stacked planar bases, which adopt chiral (helical arrangement in space) structures. CD can be used to probe DNA structures as illustrated in Fig. 17[46] since right-handed B-DNA and the left-handed Z-DNA, for example, have quite different CD spectra (not mirror images of one another since the structures are not enantiomeric). In the context of this article, the effect of DNA on the CD of bound ligands can also be a useful probe technique. Achiral molecules in solution produce no net intrinsic CD signal, however, if they bind to a chiral molecule such as DNA this can result in an induced CD (ICD) signal. Sometimes this can be used structurally but more often it gives a method of estimating binding strengths between the DNA and ligand, or the DNA base : ligand ratio. An example of this is given in Fig. 18.[47]

A useful assay developed by Jaroslav Malina (personal communication) to determine in which groove a metal complex binds is to probe the displacement of the ligand by the major groove binding $[Co(NH_3)_6]^{3+}$ (or conversely). With poly[d(G–C)]$_2$, $[Co(NH_3)_6]^{3+}$ induces Z-DNA. When it displaces a ligand and binds, the CD changes polarity at 290 nm from the positive signal for B-DNA to negative as illustrated in Fig. 19.[48] In other cases, where there is no competition between the ligand binding sites, there is no induction of Z-DNA.

3.5 Linear dichroism

Linear dichroism (LD) is a technique that may be used with systems that are either intrinsically oriented or are oriented

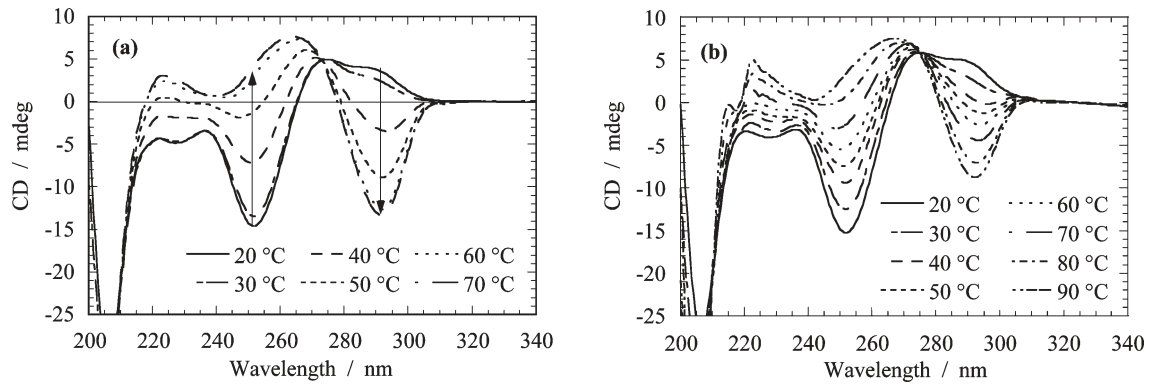


Fig. 17 CD spectra of [poly(dG–dC)]$_2$ as a function of temperature in the presence of (a) spermine (5 μM and 10 mM NaCl), 90 μM DNA and (b) $[Co(NH_3)_6]^{3+}$ (5 μM and 10 mM NaCl), 100 μM DNA. Arrows in (a) indicate direction of change in the CD spectra as the B to Z transition progresses.[46]

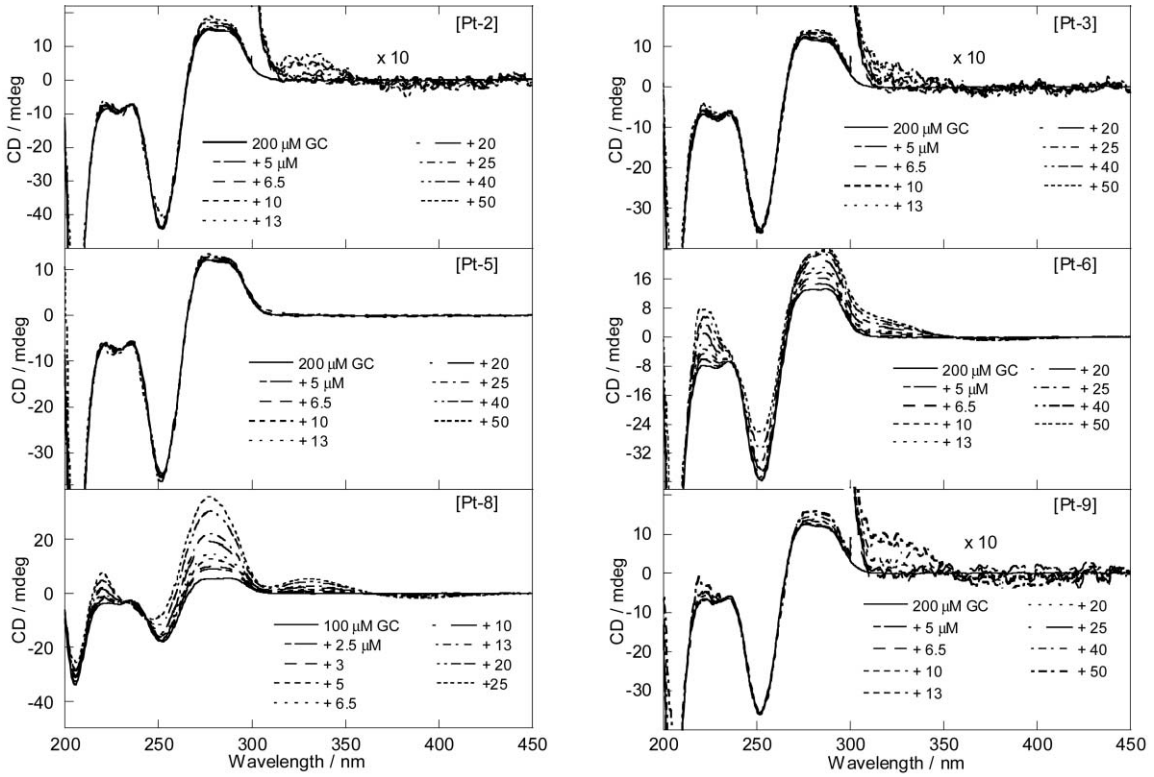


Fig. 18 CD titration series for some platinum complexes with GC DNA (200 μM, 50 mM NaCl, except for [Pt-2] where it is 20 mM NaCl and [Pt-8] where it is 100 μM DNA, 10 mM NaCl). [Pt-X] represents [Pt(L)Cl(DMSO)] (HL = R′C(O)NHC(S)NR_2: R′ = aryl, NR_2 = amine) where for [Pt-2], NR_2 = $N(CH_2CH_3)_2$ and R′ = NO_2, for [Pt-3], NR_2 = $N(CH_2CH_3)_2$ and R′ = OCH_3, for [Pt-5], NR_2 = NO4(CH_2) and R′ = NO_2, for [Pt-6], NR_2 = NO4(CH_2) and R′ = OCH_3, for [Pt-8], NR_2 = $N(CH_2CH_2OH)_2$ and R′ = NO_2 and for [Pt-9], NR_2 = $N(CH_2CH_2OH)_2$ and R′ = OCH_3. DNA base:[Pt-X] ratios are in increasing order of CD at 320 nm as indicated on the figures.[47]

during the experiment; it is the difference in absorption of light linearly polarised parallel and perpendicular to an orientation axis.[45] In the case of long DNA molecules (>250 base pairs), flow orientation such as that induced by a Couette flow cell (Fig. 20)[49] is an effective orientation technique. The speed of the revolving chamber must be such that it is not causing turbulent flow but is still strong enough to orientate the sample in question, in this case, DNA. DNA base pairs are approximately perpendicular to the DNA helix axis and their in-plane π–π* transitions give a negative LD signal under absorbance bands.

If the DNA is bent or kinked, as occurs with adding the above mentioned iron triple helicate, the DNA LD signal magnitude decreases as illustrated (Fig. 21). The bend induced per bound helicate can be determined.[29] LD can also be used to probe the binding of molecules to DNA since small molecules will only produce a flow LD signal if they are oriented due to binding in a specific regular manner to the long DNA. Fig. 21 shows an example of DNA bending (as viewed by the loss of LD at 260 nm upon addition of ligand) and also the orientation of the ligand itself in the ligand charge transfer region at 550 nm.[48]

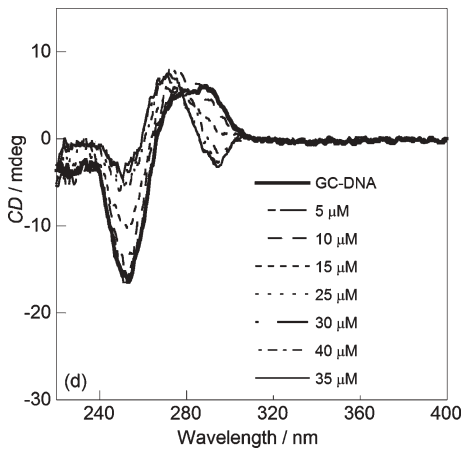


Fig. 19 Displacement from poly[d(G–C)]$_2$ (60 μM base) of a racemic methyl substituted analogue of the helicate of Fig. 2 (6 μM) by $[Co(NH_3)_6]^{3+}$ and the consequent change of the 290 nm CD signal from the positive of B-DNA to negative of Z-DNA.[48]

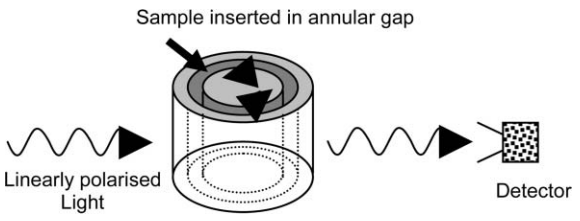


Fig. 20 Schematic diagram of Couette flow cell.[49]

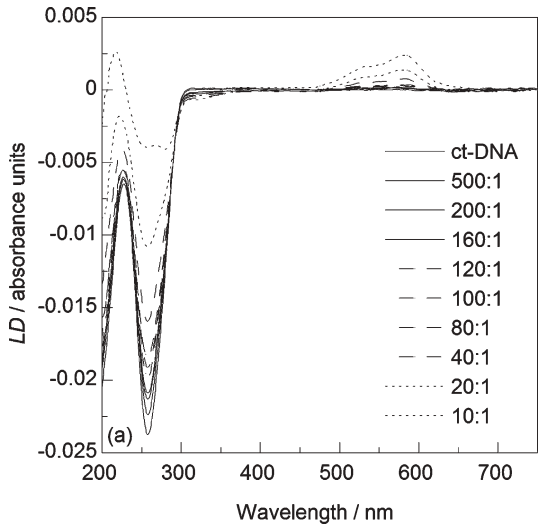


Fig. 21 LD spectra of (a) (R)–L_3, (b) (M)–L_3 and (c) (P)–L_3 binding to ct-DNA at constant DNA (500 μM) and increasing metal complex (0–50 μM) concentration. Water was subtracted from each sample spectra and the data zeroed at 420 nm for the racemic metal complex and enantiomers. L3 is the 3-methyl subsituted version of the helicate described in Fig. 2.

3.6 Molecular dynamics simulations

Molecular dynamics simulations provide an atomic-resolution depiction of molecular systems by the application of Newton's equations of motion to generate successive configurations of a system which combine to give a trajectory.[50] In biomolecular science we can obtain both physico-chemical properties through statistical sampling of a system in equilibrium and an understanding of pathways and mechanisms of action by following individual particle motions. Comparing simulation results with experimental data is essential because the experimental data can be used to validate the simulations and the simulations may help to interpret the experiments correctly or guide future work. The method is limited, however, by the necessity for an accurate description of the experimental situation and the computational power needed to accurately describe the forces and motions involved. Concessions have to be made in order to run simulations on a reasonable timescale, such as limiting the number of degrees of freedom, the length of time simulated, the system size, whether solvent molecules are explicitly included and the accuracy of the relevant force fields. Relevant controllable parameters in such experiments include the temperature, pressure, volume of periodic box, and of course the total length of the simulation. Most DNA simulations are on relatively short DNA sequences which precludes observation of any long range effects.

Khalid *et al.*[51] have demonstrated the usefulness of molecular dynamics by analysing the molecular interactions that lead to binding and subsequent DNA structure control by the metallohelicate illustrated in Fig. 10. They discovered by simulating over an array of parameters that the helicate-DNA interactions that promote DNA coiling are dominated not by the cationic charge of the molecule (as in the case of DNA condensation into toroids) but by short-range interactions.

Haworth *et al.* also found molecular modelling useful as noted above in determining the binding modes of cobalt amine complexes and understanding how these metal complexes influence DNA structural changes such as the B to Z transition and DNA condensation.[52]

All-atom molecular modelling as described above is limited by the time and computational power required to perform the simulation—particularly in relation to the volume of the system and the total simulation time. One of the ways to extend the usefulness of MD is to reduce the number of degrees of freedom in a technique known as coarse-grain modelling.[53] Here, a number of atoms are mapped onto coarse-grain beads thus eliminating fine atom–atom interaction details and yielding a simulation that requires fewer resources and runs faster than an all-atom equivalent. However, at this point in time it is very difficult to find effective merging of molecular dynamics and course graining for DNA systems.

Conclusions

The most important molecule in any cell is DNA and structural changes induced by foreign bodies such as metallomolecules are of huge significance in the fields of molecular biology and medicine. Since the double-helical structure of B-DNA was recognised in the 1950s,[3] many synthetic molecules have been shown in bind to DNA in differing binding modes promoting a variety of structural changes in the DNA. Whilst some molecules such as *cis*-platin, a leading chemotherapy agent, have been found to be DNA-targeting (a long time after its introduction into the drug market in 1978), other molecules such as an array of double and triple metallohelicates[30] have been specifically designed with DNA as the target.

The binding modes of metallomolecules can be divided into an external association with DNA, groove binding in either the major or minor groove and intercalation between DNA base pairs. The precise binding mode and consequent DNA structural changes can be analysed using a range of experimental techniques, such as NMR, X-ray crystallography, gel electrophoresis, circular dichroism, linear dichroism and modelling using molecular dynamics. Each of these techniques has different advantages and different limitations, such as the need to generate crystals for X-ray crystallography or the necessity for molecular dynamics simulations to accurately represent the *in vivo* situation. Despite these restrictions they have proved reliable and accurate techniques when used appropriately for discerning the exact nature of DNA–metallomolecule interactions.

Synthetic metallomolecules have been proven to have a range of dramatic effects on the structure of DNA. As synthetic and analytical chemistry continues to evolve, metallomolecules have a bright future both in aiding our understanding of the structure and function of DNA and also as potential diagnostic or therapeutic agents.

References

1 F. Miescher, *Med. Chem. Unters.*, 1871, **4**, 441.
2 E. Chargaff, *Experientia*, 1950, **6**, 201.
3 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964.

4 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.
5 W. Guschlbauer, *Nucleic Acid Structure*, Springer-Verlag New York Inc., New York, 1976.
6 S. Neidle, *DNA Structure and Recognition*, OUP, Oxford, 1994.
7 R. Holliday, *Genet. Res.*, 1964, **5**, 282.
8 G. L. Eichhorn and Y. A. Shin, *J. Am. Chem. Soc.*, 1968, **90**, 7323.
9 E. F. Gale, E. Cundliffe, P. E. Reynolds, M. H. Richmond and M. J. Waring, *The Molecular Basis of Antibiotic Action*, John Wiley & Sons Ltd., Chichester, 2nd edn, 1981.
10 D. Goodsell and R. E. Dicherson, *J. Med. Chem.*, 1986, **29**, 727.
11 D. Z. M. Coggan, P. J. Baters, I. S. Haworth, A. Rodger and A. J. Robinson, *Inorg. Chem.*, 1999, **38**, 4486.
12 R. Marrington, T. R. Dafforn, D. J. Halsall and A. Rodger, *Biophys. J.*, 2004, **87**, 2002.
13 B. Rosenberg, L. Van Camp and T. Krigas, *Nature*, 1965, **205**, 698.
14 T. Boulikas and M. Vougiouka, *Oncol. Rep.*, 2003, **10**, 1663 and references within.
15 J. H. J. den Hartog, C. Altona, J-C. Chottard, J-P. Girault, J-Y. Lallemand, F. A. A. M. de Leeuw, A. T. M. Marcelis and J. Reedijk, *Nucleic Acids Res.*, 1982, **10**, 4715.
16 S. E. Sherman, D. Gibson, A. H. J. Wang and S. J. Lippard, *J. Am. Chem. Soc.*, 1988, **110**, 7368.
17 Y. Zou, B. Van Houten and N. Farrell, *Biochemistry*, 1993, **32**, 9632.
18 C. Manzotti, G. Pratesi, E. Menta, R. Di Domenico, E. Cavalletti, H. H. Fiebig, L. R. Kelland, N. Farrell, D. Polizzi, R. Supino, G. Pezzoni and F. Zunino, *Clin. Cancer Res.*, 2000, **6**, 2626.
19 M. J. Clarke, F. Zhu and D. R. Frasca, *Chem. Rev.*, 1999, **99**, 2511.
20 E. Alessio, E. Iengo, S. Zorzet, A. Bergamo, M. Coluccia, A. Boccarelli and G. Sava, *J. Inorg. Chem.*, 2000, **79**, 173.
21 C. Hiort, B. Norden and A. Rodger, *J. Am. Chem. Soc.*, 1990, **112**, 1971.
22 D. M. Loprete and K. A. Hartman, *Biochemistry*, 1993, **32**, 4077.
23 H. Tabor, *Biochemistry*, 1962, **1**, 496.
24 M. H. A. Parkinson, M. Hall, K. J. Sanders and A. Rodger, *Phys. Chem. Chem. Phys.*, 2000, **2**, 5469.
25 R. V. Gessner, G. J. Quigley, A. H. J. Wang, G. A. V. D. Marcel, J. H. V. Boom and A. Rich, *Biochemistry*, 1985, **24**, 237.
26 A. Rodger, A. Parkinson and S. Best, *Eur. J. Inorg. Chem.*, 2001, **9**, 2311.
27 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, *Chem. Commun.*, 1997, **9**, 1807.
28 A. Oleksi, A. G. Blanco, R. Boer, I. Uson, J. Aymami, A. Rodger, M. J. Hannon and M. Coll, *Angew. Chem., Int. Ed.*, 2006, **45**, 1227.
29 M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chem., Int. Ed.*, 2001, **40**, 879.
30 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger and M. J. Hannon, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5069.
31 M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, J. Wiley & Sons Ltd, Chichester, 2001.
32 W. I. P. Mainwaring, J. H. Parish, J. D. Pickering and N. H. Mann, *Nucleic Acid Biochemistry and Molecular Biology*, Blackwell Scientific Publications, Oxford, 1982.
33 L. S. Lerman, *J. Mol. Biol.*, 1961, **3**, 18.
34 L. S. Lerman, *Proc. Natl. Acad. Sci. U. S. A.*, 1963, **49**, 94.
35 M. Laine, F. Richards, E. Tarnaud, C. Bied-Charreton and C. Verchere-Beaur, *J. Biol. Inorg. Chem.*, 2004, **9**, 550.
36 J. Liu, T. Zhang, T. Lu, L. Qu, H. Zhou, Q. Zhang and L. Li, *J. Inorg. Biochem.*, 2002, **91**, 269.
37 I. Haq, P. Lincoln, D. Suh, B. Norden, B. Z. Chowdhry and J. B. Chaires, *J. Am. Chem. Soc.*, 1995, **117**, 4788.
38 E. Molderheim, M. J. Hannon, I. Meistermann, A. Rodger and E. Sletten, *J. Biol. Inorg. Chem.*, 2002, **7**, 770.
39 W. L. Bragg, *Proc. Cambridge Philos. Soc.*, 1913, **17**, 43.
40 M. H. Ushay, T. D. Tullis and S. J. Lippard, *Biochemistry*, 1981, **20**, 3744.
41 U. Bierbach, Y. Qu, T. W. Hambley, J. Peroutka, H. L. Nguyen, M. Doedee and N. Farrell, *Inorg. Chem.*, 1999, **38**, 3535.
42 M. J. Browning, PhD thesis, *Estrogen steroid derived metallodrugs*, University of Warwick, 2006.
43 G. F. Carle and M. V. Olson, *Nucleic Acids Res.*, 1984, **12**, 5647.
44 J. Pelta, F. Liovolant and J-L. Sikorav, *J. Biol. Chem.*, 1996, **271**, 5656.
45 A. Rodger and B. Norden, *Circular Dichroism and Linear Dichroism*, OUP, Oxford, 1997.
46 A. Parkinson, PhD thesis, *DNA–amine interactions*, University of Warwick, 1998.
47 A. Rodger, K. K. Patel, K. J. Sanders, M. Datt, C. Sacht and M. J. Hannon, *J. Chem. Soc., Dalton Trans.*, 2002, 3656.
48 J. Peberdy, PhD thesis, *Biomacromolecule–ligand interactions*, University of Warwick, 2004.
49 W. S. Brey, *Physical Chemistry and its Biological Applications*, Academic Press Inc., Burlington, MA, 1978.
50 A. Rodger, H. C. Latham, P. Wormell, A. Parkinson, M. J. Ismail and K. J. Sanders, *Enantiomer*, 1998, **3**, 395.
51 S. Khalid, M. J. Hannon, A. Rodger and P. M. Rodger, *Chem.–Eur. J.*, 2006, **12**, 3493.
52 A. Rodger, K. J. Sanders, M. J. Hannon, I. Meistermann, A. Parkinson, D. S. Vidler and I. S. Haworth, *Chirality*, 2000, **12**, 221.
53 J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Muller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter and V. Tries, *Adv. Polym. Sci.*, 2000, **152**, 41.

Publication 117

doi:10.1016/j.jmb.2006.12.057 J. Mol. Biol. (2007) 366, 1558–1568

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Structural Characterisation of the Insecticidal Toxin XptA1, Reveals a 1.15 MDa Tetramer with a Cage-like Structure

Sarah C. Lee[1,3]*, Svetla Stoilova-Mcphie[2], Laura Baxter[1], Vilmos Fülöp[2] Janey Henderson[3], Alison Rodger[4], David I. Roper[2], David J. Scott[5] Corinne J. Smith[2] and J. Alun W. Morgan[1]

[1]*Warwick HRI, University of Warwick, Wellesbourne Warwick CV35 9EF, UK*

[2]*Department of Biological Sciences, University of Warwick, Gibbet Hill Road Coventry, CV4 7AL, UK*

[3]*Department of Biological and Molecular Sciences Faculty of Health and Life Sciences, Coventry University Priory Street, Coventry CV1 5FB, UK*

[4]*Department of Chemistry University of Warwick Gibbet Hill Road, Coventry CV4 7AL, UK*

[5]*National Centre for Macromolecular Hydrodynamics School of Biosciences University of Nottingham Sutton Bonington Leicestershire, LE12 5RD, UK*

**Corresponding author*

A recently identified class of proteins conferring insecticidal activity to several bacteria within the *Enterobacteriaceae* family have potential for control of commercially important insect pests. Here, we report the first purification, biophysical characterisation and 3-D structural analysis of one of the toxin components, XptA1, from *Xenorhabdus nematophila* PMFI296 to a resolution of 23 Å. Membrane binding studies indicate that the three-component toxin system has a different mode of action from that of proteins from *Bacillus thuringiensis (Bt)*. Biophysical characterisation of XptA1 suggests a mechanism of action of XptA1 whereby it first binds to the cell membrane forming a structure with a central cavity and forms a complex with its partners XptB1 and XptC1 producing the full insecticidal toxin. The structure of XptA1 is shown by a combination of electron microscopy, ultracentrifugation and circular dichroism spectroscopy to be a 1.15 MDa tetramer with a cage-like structure. Each of the four symmetry-related subunits has three well-defined domains and a longitudinal twist with one end narrower than the other. One third of the residues of XptA1 are α-helical and it is suggested the subunits associate partly *via* an α-helical coiled-coil interaction. XptA1 itself shows the same secondary structure at neutral pH and in an alkaline environment up to pH10.5. This pH tolerance indicates that the folded XptA1 can pass through the midgut of Lepidopteran insects susceptible to the insecticidal toxin complex. This implies therefore that its folded structure is important for its biological activity.

Keywords: insecticidal toxin; *Xenorhabdus nematophila*; *Bacillus thuringiensis*; tetramer; cage-like structure

Introduction

There is increasing demand for selective insect toxins whose mode of action differs from the currently used *Bacillus thuringiensis (Bt)* protein toxins.[1] *Bt* toxins have been used for the control of insect pests in commercial crops for more than 40 years, initially as a sprayable application, and in more recent years in transgenic crops. Finding an alternative protein insecticide that has a selective host range and different mode of action from the *Bt* toxins for the control of commercially important insect pests is therefore of great interest. In the case

Present addresses: S. C. Lee, Warwick HRI, University of Warwick, Wellesbourne, Warwick CV35 9EF, UK; J. Henderson, School of Science and Technology, University of Teesside, Middlesbrough, TS1 3BA, UK.

Abbreviations used: *Bt*, *Bacillus thuringiensis*; BBMV, brush border membrane vesicle; TEM, transmission electron microscopy; FSC, Fourier shell correlation.

E-mail address of the corresponding author: s.c.lee@warwick.ac.uk

of *Bt*, the insecticidal toxins are produced as protoxins and form a crystalline inclusion that is solubilised during passage through the insect gut.[2,3] The high pH conditions in the midgut and the action of proteases solubilise the protein and produce a smaller active toxin. By way of contrast, the class of toxins, which is the focus of this work, are initially synthesised as the final product, so it is reasonable to expect that their mechanism of action would differ from that of *Bt* toxins.

Two of the new class of toxins, which are from *Xenorhabdus* spp.[4] and *Photorhabdus luminescens*,[5,6] live symbiotically with insect parasitic nematodes. In this symbiotic relationship, the nematode worm delivers the bacteria into the insect host, the insect dies and the nematode and bacteria benefit from the resulting insect cadava. It is not clear what role the toxins play in this scenario, but it is believed that they either play a role in killing the target insect, or in helping maintain the insect cadaver as a unit for nematode and bacterial growth in the environment by preventing other insects feeding on it.[7–9] Another example is *Serratia entomophila*, a specific pathogen of the New Zealand grass grub,[10] which produces toxins that are involved in the blockage of the insect gut and aid bacterial colonization of the insect, causing septicaemia.[11,12] In these three *Enterobacteriaceae* a common family of toxin genes have been characterised and correlated to insecticidal activity.[13]

The motivation behind this study is that it has been proposed that, as with *Bt* toxins, these toxins could be used directly as a target-selective sprayable product to control insects or, alternatively, the toxin genes could be used to produce transgenic plants that are directly resistant to insects. In support of the latter strategy, the homologous TcdA1 toxin gene (*tcbD*) from *P. luminescens* has already been expressed in a plant for insect control.[14]

The potential for use of this class of naturally occurring protein toxins to protect plants and/or control insects is almost certainly not limited to the examples cited where the concept has been proven. Importantly these genes are not just restricted to these three bacteria but are now recognised in a growing number of species and therefore are emerging as a new bacterial toxin class. Genes homologous to the proteins in the above insecticidal toxin complexes have been found in *Pseudomonas syringae* pv. *tomato*,[15] *Fibrobacter succinogenes*, *Treponema denticola*,[16] *Yersinia enterocolitica*,[17] *Yersinia pestis*,[18] and *Yersinia pseudotuberculosis* (NCBI Microbial Genomes Annotation Project). In *Yersina*, the role of the gene-products is unclear; however, it has been suggested that they may be important in the association of *Y. pestis* and *Y. pseudotuberculosis* with insects and involved in transmission of the pathogen.[18] Recent work carried out in *Y. enterocolitica* shows that the proteins may facilitate the persistence of the bacterium in the gastrointestinal tract of mice[17] and also aid the survival of pathogenic strains outside mammalian hosts.[19] This work therefore has implications for the role and mechanism of action of this growing protein family that are of wide scientific, commercial and medical interest.

To date, there is no structural information available on any of the gene products (or any homologous proteins), which are believed to be the key to the insecticidal behaviour of these toxin complexes. We therefore decided to undertake an extensive study to provide such data on one of the gene products from one organism: namely XptA1 from *Xenorhabdus nematophila* to provide a foundation for future mechanistic studies on this and related systems.

Two insecticidal toxin complexes from *X. nematophila* have been described which show activity against a broad-range of insect pests including *Pieris brassicae*, *Plutella xylostella*, *Heliothis virescens* and *Heliothis zea*†.[4,19] Each complex involves three proteins (their size from primary sequence indicated in parentheses): XptA1 (287 kDa), XptB1 (110 kDa), XptC1 (158 kDa); and XptA2 (284 kDa) XptB1, XptC1. We have chosen to focus on XptA1 as the XptA proteins have been found to control host-range specificity (XptA1 directs insecticidal activity towards *P. brassicae* while XptA2 directs insecticidal activity towards *H. virescens*) and are believed to be the starting point for the assembly of the active complexes. In addition, XptA1 has been shown to be toxic to target insects by itself (though its activity is greatly increased in the presence of XptB1 and XptC1[20]), and a homologue of XptA1, TcdA1 (*P. luminescens*) has been expressed in plants and by itself shows activity towards sensitive insects.[14] Thus, we predict that XptA1 is the key to the toxicity of the XptA1/B1/C1 complex and is hence the focus here.

This paper describes the purification of XptA1, studies on its interaction with target membranes and cells, its biophysical properties and the first structural characterisation to produce a three-dimensional model at 23 Å resolution.

Results

Binding of Xpt toxin complexes to target vesicles and cells

Samples from the initial toxin complex, toxin complex–brush border membrane vesicles (BBMV) mixture and the final bound BBMV pellet, were examined by SDS-PAGE and Western blot analysis (Figure 1). In the wild-type samples, XptA1 and XptA2 can be seen in the starting material (lane 1), but only XptA1 is present in the final pellet sample (lane 2), and hence had bound specifically to the vesicles. No change in the size of XptA1 was observed on binding to BBMV. Material prepared from δxptA1 lacked the XptA1 protein (lane 3), and no protein was detected bound to the BBMV (lane 4).

† Jarrett, P. D., Ellis. D. & Morgan, J. A. W. (1997). In *World Intellectual Property*, Patent GB 97/02284.

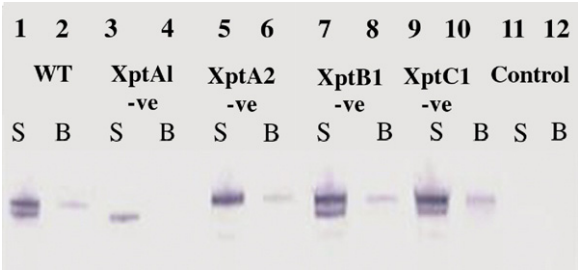


Figure 1. SDS-PAGE Western blot analysis of XptA component of the Tc toxin complex binding to *P. brassicae* BBMVs. **S**, starting Tc sample; and **B**, material bound to the BBMVs. The starting Tc sample was prepared from PMFI296 (wt), *xpt*A1:kan (XptA1-ve), *xpt*A2:kan (XptA2-ve), *xpt*B1:kan (XptB1-ve), and *xpt*C1:kan (XptC1-ve). The monoclonal antibody 2G6 binds to both XptA1 and XptA2 and was used to probe the gel.

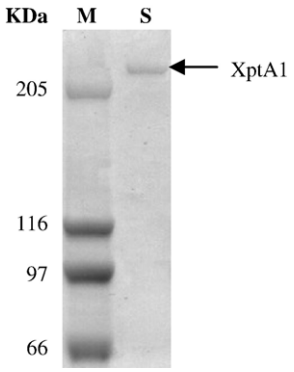


Figure 2. SDS-PAGE analysis of purified XptA1 protein, from the P_L promoter in *E. coli*. 3–8% Tris–acetate SDS-PAGE analysis of XptA1 purified by fast protein liquid chromatography gel filtration on a Superose 6 FF column. Markers (M) and their size are indicated, and the position of purified XptA1 is marked (arrow).

This again indicates that XptA2 does not bind to BBMV from *P. brassicae*. The material prepared from strain δ*xpt*A2 lacked XptA2 but contained XptA1 (lane 5), and after incubation with BBMV XptA1 was detected bound to the vesicles (lane 6). XptA1 and XptA2 were both detected in the starting material prepared from the strains δ*xpt*B1 and δ*xpt*C1 (lanes 7 and 9), but only XptA1 was detected bound to the BBMV (lanes 8 and 10). Control BBMVs were analysed and no cross-reactivity of the monoclonal antibody 2G6 to BBMV proteins was detected.

Binding experiments were also carried out using cultured insect *sf*21 cells in place of *P. brassicae* BBMVs, the experiment being carried out in an identical manner. The same pattern of protein binding was observed for cells as was seen for vesicles, in that XptA1 bound to Sf21 cells regardless of whether XptA2, XptB1 and XptC1 were present, and that XptA2 did not bind to Sf21 cells. Again the full-length XptA1 protein had bound to the cells.

Expression and purification of XptA1

The XptA1 protein was expressed at a high level in *Escherichia coli* (pLEX-*xptA1*) but most of the protein was found to be present in inclusion bodies. By altering the growth conditions and screening sonication buffers, a system for the production of soluble XptA1 was derived (as outlined in the Materials and Methods). By exploiting the large size of XptA1 protein a simple purification scheme was established that yielded large quantities of protein which ran as a single dominant band when analysed by SDS-PAGE (Figure 2). The keys to the success of this method were dialysis using a high molecular mass cutoff filter (500 kDa) and gel filtration using Superose 6 media which is designed to separate large proteins.

Secondary structure estimation of XptA1

Experimental determination of secondary structure content was made using circular dichroism spectroscopy[21] (Figure 3). The theoretical extinction coefficient per amino acid residue for XptA1 is 144 (mol amino acids)$^{-1}$ dm^3 cm^{-1} (determined from the aromatic residues in the protein sequence: XptA1-Q93RP0). Using absorption data the estimated protein concentration of the protein used to generate the data in Figure 3 is 400–500 μM in amino acid residues. Data collected at higher concentrations in 0.1 mm path length cuvettes are consistent with the pH 8 data at lower concentration leading us to conclude there are no concentration-dependent structural changes. The visual appearance of the spectra suggests they have high α-helix content with negative maxima at 208 nm and 222 nm. A more quantitative estimate of the secondary structure content of the protein was obtained from CDsstr[22] using concentrations determined from absorbance

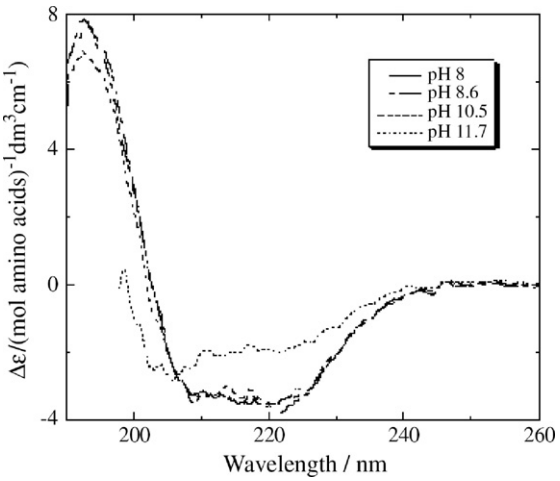


Figure 3. CD spectra of XptA1 used to predict its secondary structure and observe the effects of a change in the pH. Continuous line, pH 8; broken line, pH 8.6; small broken line, pH 10.5; dotted line pH 11.7. Protein was placed in an aqueous solution of NaOH to give the stated pH, monitored at a path length of 1 mm, scan speed of 100 nm, bandwidth 1 nm, 16 scans and a response time of 1 s.

spectroscopy. The protein has approximately 34–40% α-helix, 6% 3-10 helix, 12–14% β-sheet, 11% turns, 4–5% poly proline II and 27–30% other.

The effect of pH on secondary structure was monitored by CD spectroscopy (Figure 3). The protein was stable up to a pH of 10.5 with little change in the CD spectra. At a pH of pH11.5 there was a noticeable loss of α-helical content, and an increase in random coil as indicated by a shift of the 208 nm (α-helix) peak toward the 200 nm position (random coil). CDsstr indicates an increased poly-proline type II content in accord with this since random coil and poly-proline II both have negative bands at 200 nm and folded protein data bases do not include random coil motifs.[22]

Analytical ultracentrifugation and dynamic light scattering analysis of XptA1

The oligomeric state of the XptA1 protein was studied using analytical ultracentrifugation[23–25] and dynamic light scattering. Sedimentation equilibrium experiments indicated a molecular mass of 1.1 MDa (Figure 4), consistent with a tetramer (1.148 MDa based upon sequence information corrected to 1.15 MDa). Analysis of XptA1 by dynamic light scattering indicated that the protein was again predominately monodispersed. The estimated particle size was 1.1 MDa, with an estimated spherical diameter of 20 nm. These results suggested that the XptA1 protein was large enough for its gross ultra structure to be observed by electron microscopy.

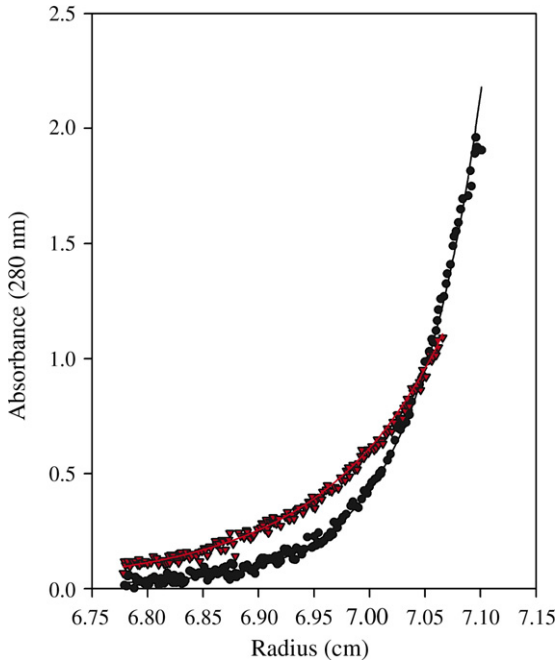


Figure 4. Analytical ultracentrifugation data showing the molecular mass of the purified XptA1 protein. Typical sedimentation equilibrium data of XptA1 at rotor speeds of 12,000 rpm (red) and 16,000 rpm (black). The data were fitted using the program SEDFIT[33] to a single species model. The protein was shown to have a molecular mass of 1.15 MDa, consistent with the protein being a tetramer.

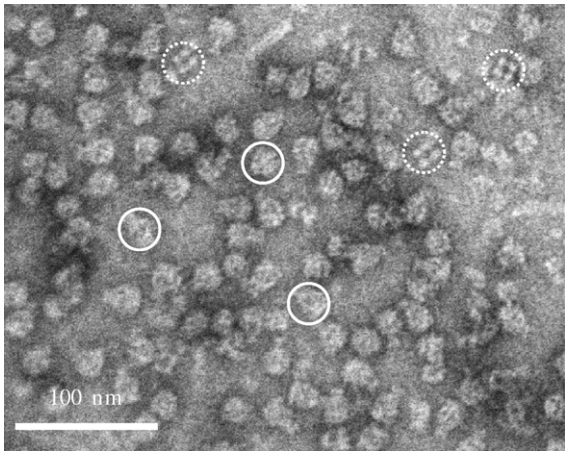


Figure 5. An example of negatively stained XptA1 particles from a digital micrograph. Pure XptA1 protein stained with 2% uranyl acetate. Continuous and dotted circles indicate different views of XptA1 suggesting an overall bottle-like structure composed of four subunits. The scale bar represents 100 nm.

Electron microscopy of XptA1

The XptA1 protein was successfully viewed by transmission electron microscopy (TEM) using uranyl acetate as a negative stain. The micrographs (Figure 5) revealed two characteristic particle views: a side-on view with a "bottle-like" shape, and an end-on view with a square shape containing four separate density areas in each corner and a central dimple. A total of 18,351 individual particles were collected from the TEM images, and the particles were aligned and classified using the EMAN software. A final set of 9861 particles was included in the 3-D reconstruction. All particles appeared to be structurally related as judged by their dimensions and general features. Twenty-five example projections from the final 3-D reconstruction, spaced at 30° intervals, are presented in Figure 6. The appearance of the individual class averages and especially the top and bottom views indicated that the structure might have *C*4 symmetry. In order to test this, *C*1, *C*2, *C*3, *C*4 and *C*5 symmetry was imposed on the data set. The highest resolution and best convergence between class averages and projections was obtained by imposing *C*4 symmetry. The resolution of the final 3-D structure as estimated from the Fourier shell correlation (FSC) value = 0.5 was 23 Å. and that for which FSC = 0.3 was 20 Å. (Figure 7). The threshold corresponding to 1.2 MDa of a XptA1 tetramer was calculated to be 2.7 sigma assuming an average protein density of 1.35 Da/A.[3] The 3-D structure of XptA1 at 20 Å resolution is shown in Figure 8. The oligomeric organisation of the XptA1 complex is a tetramer with an overall bottle-like structure of 16 nm × 19 nm, with a central double ventricular chamber, consistent with the observations of the frictional ratio from sedimentation velocity. The four elongated XptA1 subunits (8 nm × 19 nm each) are held together in a supra-

Figure 6. Projections from the XptA1 3-D structure. Twenty-five representative projections spaced at 30° calculated from the final XptA1 3-D structure. The boxes size is 120×120 pixels at 2.87 Å/pixel.

molecular ring with one end narrower than the other, but both open forming a channel. There was an observed twist in the subunits down their length to generate the narrower end of the structure (Figure 8(c) and (d)).

The 3-D density map indicated that there are three distinct protein density areas well defined within each subunit (protein monomers) (Figure 8(f), i, ii and iii) suggesting there are three large domains in each subunit. Longitudinal views through the structure show interactions between the subunit at three distinct places on each one which appear to hold the structure together (Figure 8(d)). This gives the molecule an overall cage-like appearance. A large central chamber was detected, extending to the full length of the molecule and openings at either end and between each of the subunits. The overall dimensions of the cavity are 13 nm×17 nm in the middle, narrowing to 7 nm towards the top and bottom part of the structure to (Figure 8(e)).

Bioinformatic analysis of the amino acid sequence of XptA1

Secondary structure content of the protein predicted from the primary sequence (XptA1-Swiss-Prot/TrEMBL No. Q93RP0) shows that the major structural elements appear to be α-helices and coils with only a small proportion of β-sheet (13% as given by the secondary structure prediction programme, PSIPRED). Comparison with other prediction methods gave broad agreement with this result (data not shown). This supports the experimental data obtained from CD experiments. An area predicted to contain heptad repeat motif characteristic of a possible α-helical coiled-coil was identified towards the C terminus of the protein at residues 2191–2221. Such a motif may indicate a region of the protein likely to form a coiled-coil interaction with a neighbouring subunit. Heptad repeats are characterised by the first and forth amino acid being predominantly hydrophobic, and the fifth and seventh being predominantly charged or polar. This motif was also identified in XptA2, and other "A" components of the Tc family including TcdA, TcbA from *P.luminescens* and SepA from *S. entomophila.*

When the primary sequence was analysed for predicted helices and hydrophobicity, three distinct domains became evident. The first domain (A) was composed primarily of helices (residues 1–928), the second (B) was also composed of helices which are predominantly hydrophobic in nature (residues 929–1773), and the third domain (C) was also primarily helical (residues 1774–2523). The three electron-dense domains (i, ii, iii) of Figure 8(f) may

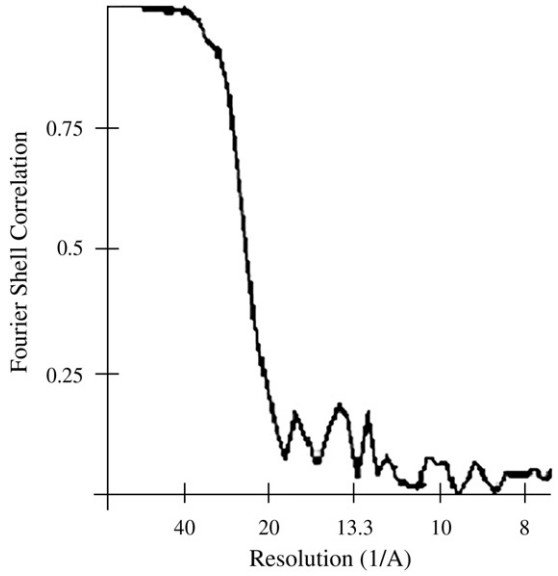


Figure 7. The Fourier shell correlation (FSC) *versus* resolution plot for the 3-D model generated with *C*4 symmetry. The FSC plot between two independent subsets of particles was obtained by splitting all the data into two halves and calculating two separate 3-D structures. The Fourier shell correlation value of 0.5 corresponds to 23 Å resolution and a value of 0.3 corresponds to 20 Å.

reflect the domains A, B and C identified through sequence analysis.

Discussion

Using a combination of biophysical techniques we have built an overall model of the macromolecular structure of the XptA1 component of the insecticidal toxin complex from *X. nematophila*. Previous work has shown that XptA1 can be added to the other components of the complex XptB1/XptC1, to give full insecticidal activity. Therefore, this assembled structure must interact with the other two components of the complex. Our initial studies indicated that XptA1 could fold, assemble into its quaternary structure, and then interact with both target vesicles and cells without any requirement for the presence of the other two toxin components. It is thus important to understand the structure of XptA1.

Both experimental CD spectroscopy and secondary structure predictions show that XptA1 is a mixed structure protein with a preponderance of α-helical residues (~35%). The remarkable stability of the secondary structure of XptA1 at pH values as high as in an insect gut (but not much higher) suggests that the observed structure is a key to its mode of action, since the toxin can pass intact through this environment before interacting with the target cell. This is in distinct contrast to *B. thuringiensis* insecticidal toxins where the proteins are produced as pro-toxins and form a crystalline inclusion that are solubilised and cleaved during passage through the insect gut. In contrast, the XptA1 protein remains intact at a high pH and the full-length protein interacts with target membranes. No processing is required prior to this interaction. We have therefore demonstrated that the overall mode of action of the *Xenorhabdus* toxin complex is quite different from that of *B. thuringiensis* toxins.

The 3-D structure of the XptA1 monomers shows it to have three sub-domains interacting *via* quite small surfaces. Database analysis of the primary sequence also indicated three sub-domains. It is possible that the sub-domains identified by data analysis correlate with those visible in the 3-D structure. Database analysis also suggests that one of the sub-domains contains a sequence, which has a propensity to form α-helical coiled-coils. This proposed coiled-coil domain is an ideal candidate for providing one of the interactions between the sub-domains.

The electron microscopy data suggest that XptA1 is a large bottle-shaped tetrameric complex with a *C*4 symmetry. The XptA1 protein has also been crystallised and initial diffraction data obtained to a resolution of 4 Å by X-ray crystallography (results not shown). These crystals are in a tetragonal *I*4 space group with one molecule in the crystallographic asymmetric unit, supporting the *C*4 symmetry observed from the EM single particle reconstruction. The proposed tetrameric structure and its overall size is also supported by analytical ultracentrifugation and dynamic light scattering results which indicate that unit size of 1.15 MDa is predominant, with only a very small percentage of molecules exhibiting other sizes. This gives us the confidence to conclude that the electron microscopy results are relevant to the native oligomeric organisation of the protein.

The large central channel open at both ends in the 3-D structure of XptA1 raises questions as to the role such a channel would play in the mechanism of toxicity. For example, is this channel involved in pore formation which is a common feature of the mode of action of many bacteria toxins, including *Bt* toxins[26,27] and pneumolysin.[28–30] In these toxins, soluble monomers are released and oligomerise at or near the target cell phospholipid membrane, forming pores, resulting in cell lysis and cell death. In contrast we demonstrate here that the XptA1 protein is assembled in solution as a tetramer, and no monomeric state is seen. In addition preliminary studies comparing the native toxin complex from *X. nematophila* PMFI296 to *Bt* Cry toxins showed that pore formation was not observed (data not shown). An alternative role for this central channel could be that XptA1 is an extrinsic membrane protein, which is able to accommodate wholly, or partially the other toxin components in its cavity. TEM micrographs of the toxin complex from *P. luminescens*[13] show a molecule with a rounded head and protruding tail, suggesting that XptB1 and

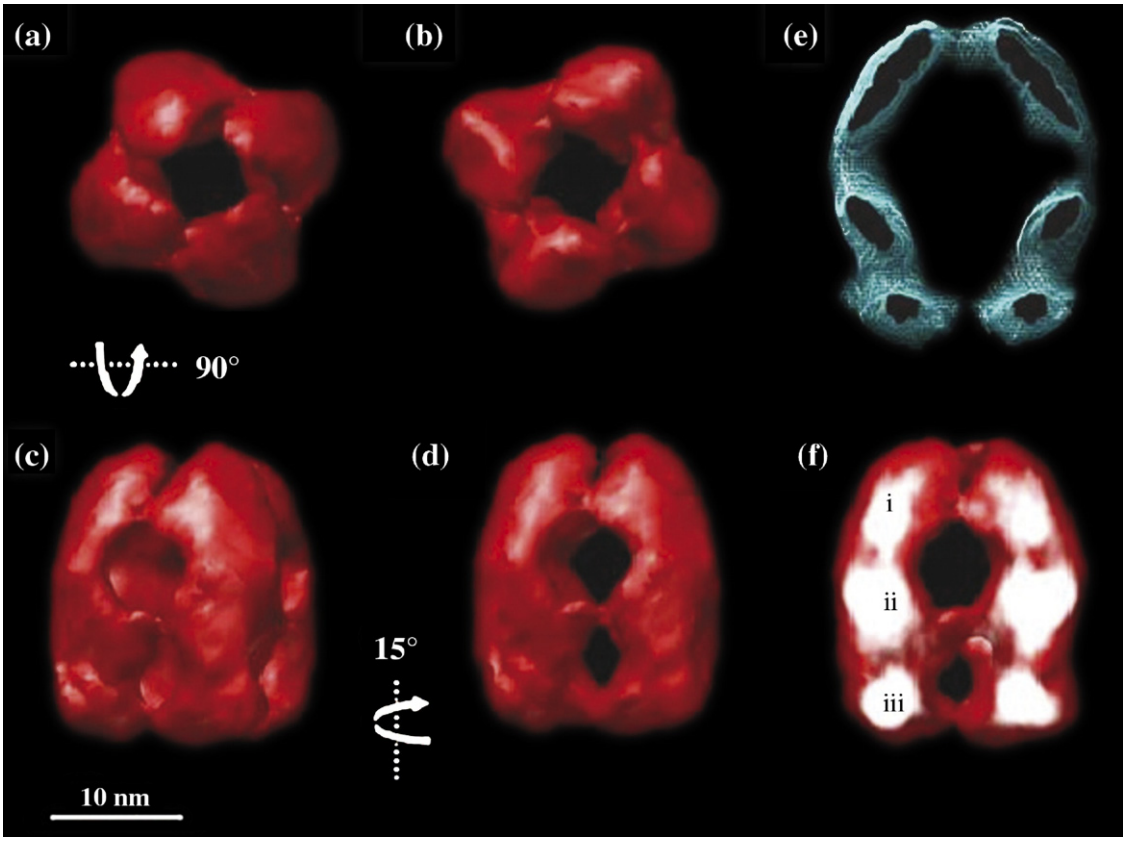


Figure 8. Three-dimensional structure of the XptA1 component of Tc insecticidal toxin complex. Surface rendered view of the (a) top; (b) bottom; (c) side view (90° rotated around the *x*-axis); (d) side view (15° rotated around the *y*-axis); (e) cross sectional view through the middle along the *y* -axis of the XptA1 complex showing the central activity chamber; (f) side view equivalent to (d) with the maximum protein density in the 3-D structure presented in white illustrating three separate electron densities for each monomer (i, ii and iii). The longitudinal length of the particle is 19 nm and width 16 nm ((c) and (d)). The four XptA1 subunits are held together in a supramolecular ring with one end narrower than the other, open-ended to form a channel. The red *iso*-surface is displayed at a threshold of 2.7 sigma, which corresponds to a 1.2 MDa complex.

XptC1 compose, at least in part, the protruding tail and may attach to XptA1 *via* its central cavity.

Conclusion

The protein XptA1 is a member of a growing family of proteins found in a number of bacteria, some of which have been shown to have insecticidal properties. To determine the detailed mechanism of the toxic effect so as to be able to develop insecticides for crop protection in the future, there is a need to establish the structure of the protein complexes involved. Here, we have undertaken the first stage of this process showing that XptA1 forms a bottle-shaped tetrameric complex whose channel does not appear to form a membrane pore but may be a binding site for one or more other proteins of the toxin complex either before or after it has bound to the target cell membrane. Although XptA1 is toxic in its own right, the increased insecticidal activity seen in the presence of XptB1 and XptC1[20] makes it important to understand where they are located.

How the complex actually brings about insect death remains to be elucidated but it is now clear that XptA1 and its complete toxin complex is unlike that of *B. thuringiensis* toxins. This is particularly attractive in the context of controlling insects resistant to *B. thuringiensis* toxins. Understanding the mechanism of action of toxins is also important as an aid to designing strategies to prevent the development of resistant insects by, for example, stacking toxin genes.[31] Given the common overall sequence size, native protein size and sequence similarity shared with XptA2,TcdA TcbA, SepA and other members of this family, the structure determined here may well be representative of these other proteins. Studies on the structure of the toxin complex or parts of it, will also help to provide a route to commercialisation of the toxins and help provide alternatives to *B. thuringiensis* products for the control of insects.

In the case of the Xpt complex, we propose that a folded bottle-shaped tetramer of XptA1 binds to the cell membrane and that it binds XptB1 and XptC1 possibly within its central channel. This assembled structure must then exert its toxic effect once

internalised within the cell. It is not clear whether the whole toxin complex XptA1/B1/C1 is delivered into the cell or if the individual components enter, leaving the XptA on the cell surface.

The results presented here for XptA1 are relevant to the study of the modes of actions of a range of homologous proteins found in a large number of species. Both further structural and mode of action analysis on these groups of toxins will help in understanding their role in invertebrate–bacterial interactions.

Materials and Methods

Bacterial strains

The wild-type strain *X. nematophila* PMFI296 and the expression strain *E. coli* (pLEX-*xpt*A1) have been described.[4] *X. nematophila* PMFI296 insertional inactivants *xpt*A1: :kan, *xpt*A2: :kan, *xpt*B1: :kan and *xpt*C1: :kan[20] were constructed by the insertion of a kanamycin resistance gene into the target toxin gene by homologous recombination. Southern hybridization, insecticidal activity tests and complementation were used to confirm gene disruption in each of the mutants.

Binding of Xpt toxins to brush border membrane vesicles and sf21 insect cells

X. nematophila PMFI296 wild-type, *xpt*A1: :kan, *xpt*A2: :kan, *xpt*B1: :kan and *xpt*C1: :kan were grown in 1 litre of Luria Broth (LB) at 30 °C with shaking (125 rpm) for two days. Cells were pelleted by centrifugation at 5000*g* for 10 min, re-suspended in 20 ml of phosphate buffered (50 mM) saline (12. 5 mM) (pH 7.2) (PBS), and sonicated eight times at 18 μ peak-to-peak for 30 s blasts, resting on ice for 30 s in between. Samples were centrifuged at 13,000*g* for 10 min and the supernatant filtered through a 0.45 μm pore-sized syringe filter. Samples were centrifuged at 100,000*g* for 60 min and the final pellet re-suspended in a small volume of PBS and filtered through a 0.2 μm –pore-size syringe filter. The protein concentration was determined using the BCA kit (Pierce, UK), and standardised to 20 mg ml^{-1} by addition of PBS.

Pieris brassicae larvae were reared on an artificial diet containing antibiotics at 18 °C as described.[4] Anterior and posterior regions of larvae were removed, and the midsections immersed in ice-cold sucrose dissection buffer (250 mM sucrose, 5 mM EDTA, 17 mM Tris (pH 7.5)), cleared of gut contents and stripped of malpighian tubules.[32] Isolated mid-guts were washed several times and stored frozen in liquid nitrogen. Samples were homogenised using a glass 100 ml tissue grinder, an equal volume of 24 mM $MgCl_2$ was added and left to stand on ice for 15 min. Gut material was centrifuged at 4500*g* for 15 min and the supernatant retained. The pellet was re-suspended in dissection buffer with an equal volume of 24 mM $MgCl_2$ and left to stand on ice for 15 min before centrifuging at 4500*g* for 15 min. The supernatants were pooled and centrifuged at 16,000*g* for 30 min. The resultant pellet was re-suspended in 0.5 volume of dissection buffer containing an equal volume of 24 mM $MgCl_2$, incubated on ice for 15 min, then centrifuged at 4500*g* for 15 min. The supernatant was centrifuged at 16,000*g* for 30 min, and the pellet consisting of BBMVs were saved. The pellets were re-suspended in PBS to a final protein concentration of 2 mg ml^{-1}, mixed with 50 μl of toxin (20 mg ml^{-1} total protein concentration) and incubated at room temperature for 1 h without shaking. The mixture was centrifuged at 13,000*g* for 15 min, washed three times by centrifugation with ice-cold PBS, and the final pellet resuspended in 200 μl of PBS. Subsamples of the starting material (unbound) and final vesicle pellet (bound) were analysed by SDS-PAGE and Western blotting using standard methods. The monoclonal antibody 2G6 was used and was prepared from BalbC mice injected with purified XptA1 protein (100 μg ml^{-1}). The antibody was selected during cloning using enzyme-linked immunosorbent assay (ELISA) with XptA1 bound to the microtitre plate. Once isolated the antibody was tested for reactivity with XptA1, XptA2, XptB1 and XptC1 expressed in *E. coli* using a standard Western blot. Specific reactions with both XptA1 and XptA2 were detected for this antibody.

Purification of XptA1

E. coli (pLEX-*xpt*A1) was cultured in 50 ml RMG media (Invitrogen) containing ampicillin (50 μg ml^{-1}) and incubated for 18 h at 30 °C with shaking (150 rpm). Aliquots (2 ml) were used to inoculate 22 cm^2 (Genetix) culture plates containing 250 ml of LB agar. Usually 16 plates were inoculated at a time and incubated for 24 h at 30 °C. Bacterial growth was scraped from the plates, re-suspended in 10 mM Tris (pH 8.6) 100 mM NaCl, 1 mM EDTA and centrifuged at 6000*g* for 10 min. The supernatant was removed and the pellet re-suspended in 30 ml of 100 mM Tris, 100 mM KCl 25%(v/v) glycerol, 0.1%(w/v) deoxycholate (pH 7.6). Samples were sonicated at 18 μ peak-to-peak for ten 30 s bursts with 30 s on ice in between bursts. Insoluble material was removed by centrifugation at 100,000*g* for 2 h and the supernatant filter sterilized (0.2 μm pore-size units). The sample was dialyzed in a 500 kDa molecular mass cut off PVDF membrane (Spectra Labs) against 50 mM Tris-glycine (pH 8.3) for 14 days. A 5 ml aliquot of the protein sample was concentrated using a Centricon plus 20 (Millipore) unit by centrifugation at 4000*g* for 15 min at 4 °C to give a final volume of 200 μl. The sample (ca 10 mg ml^{-1} protein) was passed down a Superpose 6 fast-flow gel filtration column in 100 mM Tris buffer (pH 8.0) containing 100 mM NaCl. The column (25 mm diameter by 320 mm) was run at a flow rate of 0.5 ml min^{-1}. Fractions (1 ml) were collected and 20 μl analysed by SDS-PAGE for the presence of the distinctive XptA1 protein (287 kDa) and immunoblotting. Fractions containing pure XptA1 were selected, pooled and concentrated using the Centricon plus 20 unit and re-applied to the Superose 6 fast-flow column as described. Fractions containing pure XptA1 from five initial sample runs passed twice through the fast protein liquid chromatogrphy system were pooled and concentrated using a Centricon plus 20 unit. Protein estimations were carried out using the commercial BCA method and also using absorbance at 280 nm. Samples were diluted to 10 mg ml^{-1} protein and stored at 4 °C.

Circular dichroism (CD) spectroscopy

Circular dichroism spectra were obtained on a Jasco (Tokyo, Japan) J-715 spectropolarimeter run at room temperature using a 0.1 mm path length quartz cuvette for an ~0.5 mg ml^{-1} protein solution in phosphate buffer (pH 8.0), 50 mM NaCl and a 1 mm pathlength cuvette for 0.05 mg ml^{-1} protein solutions at a range of pH values.

To monitor the effect of pH on secondary structure the pH was adjusted by placing the protein in 0.1 mM to 0.0001 mM NaOH. The pH was measured after data collection using a pH meter. Data were collected every 0.2 nm from 260 nm to 190 nm, with a scanning speed of 100 nm min^{-1} and a response time of 1 s. Protein concentrations were determined from absorbance at 280 nm with data collected using a nano-drop spectrophotometer. Data were filtered to 1 nm data intervals for use in the protein structure-fitting program CDsstr.[21]

Analytical ultracentrifugation

Sedimentation equilibrium determinations were carried out in cells with six channel centrepieces with loading at 0.5, 0.25, and 0.13 mg ml^{-1}; the set temperature on the centrifuge was 20 °C. Absorbance at 280 nm of each cell was recorded and data were obtained at 12,000, 16,000, and 22,000 rpm. Molecular mass was determined using SEDFIT.[33] The solvent density was measured directly with a Anton Paar DMA 5000 density meter.

Dynamic light scattering

XptA1 samples at a concentration of 100 μg ml^{-1} at a range of pH values between pH7 and pH10.7 were loaded into a 12 μl quartz cuvette. Data acquisition was carried out using a Dynapro instrument and dynamics V6 software (Wyatt laboratories). The acquisition time was set to 10 s and 20 readings were taken per sample.

Negative staining

Samples at a protein concentration of 1 mg ml^{-1} were applied to freshly glow-discharged carbon coated copper grids and stained with 2% (w/v) uranyl acetate. Data were collected with a JEOL 2010 transmission electron microscope operating at 200 kV at a nominal magnification of 52,000× equipped with a LaB6 cathode. Images were digitally recorded on a 2048×2048 pixel slow-scan CCD camera with 15 μm per pixel resolution (Gatan Inc.). A total of 100 micrographs were recorded.

Image processing

Image processing was carried out using the programme EMAN‡.[34] A total of 18,351 particles using a box size of 120×120 pixels, (0.29 nm/program) ***were selected from 100 micrographs using the EMAN program,*** *boxer*. 9861 of those particles that exhibited a consistent size and were evenly stained were selected visually for further analysis.

The 3-D reconstruction was performed using EMAN.[34] The initial 3-D model was generated from reference-free class averages aligned by the cross-common line method with the programs *startnrclasses* and *startAny*, imposing *C*4 symmetry. Twenty rounds of refinement were carried out in three cycles (8+12+8) increasing the low-pass filter of the images from 2.5 to 2.0 and 1.4 nm and decreasing the increment generating the projections for the consecutive classification of the particles from 25° to 18 and 15. The classkeep option was set to 0.4 in the final cycle, allowing only particles well matching the class average to be included in the next refinement. The accuracy of the refinement cycles was followed by monitoring convergence of the Fourier shell correlation (FSC) coefficients of consecutive reconstructions plotted against the modulus of the resolution.

The appearance of the particles together with molecular mass information suggested that the 3-D structure may adopt *C*4 symmetry. We tested the effect of imposing *C*1, *C*2, *C*3, *C*4 and *C*5 symmetry on the final reconstruction and the convergence plots and also monitored the degree to which each particle matched the final reconstruction. *C*4 symmetry resulted in the highest quality reconstruction and the highest degree of agreement between the projections and the original particles.

The final 3-D structure was visualised using the UCSF Chimera software§.[35]

Bioinformatic analysis of the amino acid sequence of XptA1

The amino acid sequences of XptA1 and its homologues were investigated, using the following bioinformatic tools using Swiss-Prot/TrEMBL accession numbers: XptA1-Q93RP0, XptA2-93RN7, TcdA1-Q9RN43, TcbA1-O85160, SepA-Q9F9Z3. Hydropathy was plotted using the method of Kyte & Doolittle[36] with a window size of 19. Secondary structure predictions were carried out using nnpredict[37,38] and Psipred.[39–41] Domain architecture was investigated using Prodom[42] and Pfam.[43] Finally, α-helical coiled-coil regions were predicted using COILs[44–46] and Pair coil.[47]

Acknowledgements

We thank the support of the BBSRC and EPSRC for funding this work and The Wellcome Trust for support to the imaging suite at the University of Warwick. We are grateful for access and user support and the synchrotron facilities of ESRF, Grenoble, France and to Eve Shaw and Martin Sergeant at Warwick HRI for their technical assistance.

References

1. De Maagd, R. A., Bravo, A., Berry, C., Crickmore, N. & Schnepf, H. E. (2003). Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. *Annu. Rev. Genet.* **7**, 409–433.
2. Chestukhina, G. G., Kostina, L. I., Mikhailova, A. L., Tyurin, S. A., Klepikova, F. S. & Stepanov, V. M. (1982). The main features of *Bacillus thuringiensis* delta-endotoxin molecular structure. *Arch. Microbiol.* **132**, 159–162.
3. Choma, C. T. & Kaplan, H. (1990). Folding and unfolding of the protoxin from *Bacillus thuringiensis*: evidence that the toxic moiety is present in an active conformation. *Biochemistry*, **29**, 10971–10977.
4. Morgan, J. A. W., Sergeant, M., Ellis, D., Ousley, M. & Jarrett, P. D. (2001). Sequence analysis of insecticidal genes from *Xenorhabdus nematophilus* PMFI296. *Appl. Environ. Microbiol.* **67**, 2062–2069.
5. Bowen, D. J. & Ensign, J. C. (1998). Purification and

‡ http://ncmi.bcm.tmc.edu/stevel/EMAN/doc

§ http://www.cgl.ucsf.edu/chimera

characterization of a high-molecular-weight insecticidal protein complex produced by the entomopathogenic bacterium *Photorhabdus luminescens*. *Appl. Environ. Microbiol.* **64**, 3029–3035.
6. Guo, L., Fatig, R., III, Orr, G. L., Schafer, B. W., Strickland, J. A., Sukhapinda, K. *et al.* (1999). *Photorhabdus luminescens* W-14 insecticidal activity consists of at least two similar but distinct proteins—Purification and characterization of toxin A and toxin B. *J. Biol. Chem.* **274**, 9836–9842.
7. Akhurst, R. J. (1983). Taxonomic study of *Xenorhabdus*, a genus of bacteria. Symbiotically associated with insect-pathogenic nematodes. *Int. J. Syst. Bacteriol.* **33**, 38–45.
8. Akhurst, R. J. & Boemare, N. E. (1988). A numerical study of the genus *Xenorhabdus* (*Enterobacteriaceae*) and proposed elevation of the subspecies of *Xenorhabdus nematophilus* to species. *J. Gen. Microbiol.* **134**, 1835–1845.
9. ffrench-Constant, R. H. & Waterfield, N. (2006). Ground control for insect pests. *Nature Biotechnol.* **24**, 660–661.
10. Jackson, T. A., Huger, A. M. & Glare, T. R. (1993). Pathology of amber disease in the New Zealand grass grub *Costelytra zealandica* (Coleoptera: Scarabaeidae). *J. Invertebr. Path.* **61**, 123–130.
11. Hurst, M. R. H., Glare, T. R., Jackson, T. A. & Ronson, C. W. (2000). Plasmid-located pathogenicity determinants of *Serratia entomophila*, the causal agent of amber disease of grass grub, show similarity to the insecticidal toxins of *Photorhabdus luminescens*. *J. Bacteriol.* **182**, 5127–5138.
12. Hurst, M. R. H., Glare, T. R. & Jackson, T. A. (2004). Cloning *Serratia entomophila* antifeeding genes–a putative defective prophage active against the grass grub *Costelytra zealandica*. *J. Bacteriol.* **186**, 5116–5128.
13. Waterfield, N., Bowen, R. D. J., Fetherston, J. D., Perry, R. D. & ffrench-Constant, R. H. (2001). The toxin complex genes of *Photorhabdus*: a growing gene family. *Trends Microbiol.* **9**, 185–191.
14. Liu, D., Burton, S., Glancy, T., Li, Z. S., Hampton, R., Meade, T. & Merlo, D. J. (2003). Insect resistance conferred by 283-kDa *Photorhabdus luminescens* protein TcdA in *Arabidopsis thaliana*. *Nature Biotechnol.* **10**, 1222–1228.
15. Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L. *et al.* (2003). The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. *Proc. Natl Acad. Sci. USA*, **100**, 10181–10186.
16. ffrench-Constant, R. H., Waterfield, N. R., Daborn, P., Joyce, S., Bennett, H., Au, C. *et al.* (2003). Photorhabdus: towards a functional genomic analysis of a symbiont and pathogen. *FEMS Microbiol. Rev.* **26**, 433–456.
17. Tennant, S. M., Skinner, N. A., Joe, A. & Robinson-Browne, R. M. (2005). Homologues of insecticidal toxin complex genes in *Yersinia enterocolitica* biotype 1A and their contribution to virulence. *Infect. Immun.* **73**, 6860–6867.
18. Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D., Wain, J. *et al.* (2001). Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature*, **413**, 848–852.
19. Bresolin, G., Morgan, J. A. W., Ilgen, D., Scherer, S., Scherer, F. & Thilo, M. (2006). Low temperature-induced insecticidal activity of *Yersinia enterocolitica*. *Mol. Microbiol.* **59**, 503–512.
20. Sergeant, M., Jarrett, P. D., Ousley, M. & Morgan, J. A. W. (2003). Interactions of insecticidal toxin gene products from *Xenorhabdus nematophilus* PMFI296. *Appl. Environl. Microbiol.* **69**, 3344–3349.
21. Rodger, A. & Nordén, B. (1997). *Circular Dichroism and Linear Dichroism.* Oxford University Press, OUP Oxford, UK.
22. Johnson, W. C. (1999). Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins: Struct. Funct. Genet.* **35**, 307–312.
23. Hensley, P. (1996). Defining the structure and stability of macromoleculer assemblies in solution: the re-emergence of analytical ultracentrifugation as a practical tool. *Structure*, **4**, 367–373.
24. Laue, T. M. & Stafford, W. F., III (1999). Modern applications of analytical ultracentrifugation. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 75–138.
25. Demeler, B., Saber, H. & Hansen, J. C. (1997). Identification and interpretation of complexity in sedimentation velocity boundaries. *Biophys. J.* **72**, 397–407.
26. Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J. *et al.* (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol.* **62**, 775–806.
27. Whalon, M. E. & Wingerd, B. A. (2003). Bt: mode of action and use. *Arch. Insect Biochem. Physiol.* **54**, 200–211.
28. Alouf, J. E. (2000). Cholesterol-binding cytolytic protein toxins. *Int. J. Med. Microbiol.* **290**, 351–356.
29. Gilbert, R. J. C. (2002). Pore-forming toxins. *Cell. Mol. Life Sci.* **59**, 832–844.
30. Tilley, S., Orlova, E., Gilbert, R., Andrew, P. & Saibil, H. (2005). Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell*, **121**, 247–256.
31. Waterfield, N., Hares, M., Yang, G., Dowling, A. & ffrench-Constant, R. H. (2005). Potentiation and cellular phenotypes of the insecticidal toxin complexes of *Photorhabdus bacteria*. *Cell Microbiol.* **7**, 373–382.
32. Wolfersberger, M. G., Lüthy, P., Maurer, A., Parenti, P., Sacchi, V. F., Giordana, B. & Hanozet, G. M. (1987). Preparation and partial characterization of amino acid transporting brush border membrane vesicles from the larval midgut of the cabbage butterfly (*Pieris brassicae*). *Comp. Bioch. Physiol. ser. A*, **86**, 301–308.
33. Schuck, P. (1998). Sedimentation analysis of non-interacting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* **75**, 1503–1512.
34. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semi automated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97.
35. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
36. Kyte, J. & Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132.
37. McClelland, J. L. & Rumelhart, D. E. (1988). *Explorations in Parallel Distributed Processing*, vol. 3, pp. 318–362. MIT Press, Cambridge, MA.
38. Kneller, D. G., Cohen, F. E. & Langridge, R. (1990). Improvements in protein secondary structure prediction by an enhanced neural network. *J. Mol. Biol.* **21**, 71–182.

39. Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S. & Jones, D. T. (2005). Protein structure prediction servers at University College London. *Nucl. Acids Res.* **33**(Web Server issue), W36–W38.
40. McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404–405.
41. Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202.
42. Bru, C., Courcelle, E., Carrère, S., Beausse, Y., Dalmar, S. & Kahn, D. (2005). The ProDom database of protein domain families: more emphasis on 3D. *Nucl. Acids Res.* **33**, 212–215.
43. Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T. *et al.* (2006). Pfam: clans, web tools and services. *Nucl. Acids Res.* **34**, 247–251.
44. Lupas, A., Van Dyke, M. & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
45. Lupas, A. (1996). Prediction and analysis of coiled-coil structure. *Struct. Methods Enzymol.* **266**, 513–525.
46. Parry, D. A. D. (1982). Coiled-coils in alpha-helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. *Biosci. Rep.* **2**, 1017–1024.
47. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M. & Kim, P. S. (1995). Predicting coiled coils by use of pairwise residue correlations. *Proc. Natl Acad. Sci. USA*, **92**, 8259–8263.

Edited by R. Huber

(Received 19 September 2006; received in revised form 16 December 2006; accepted 20 December 2006)
Available online 23 December 2006

PAPER www.rsc.org/dalton | Dalton Transactions

Enantiomeric resolution of supramolecular helicates with different surface topographies†

Jessica M. C. A. Kerckhoffs,[a] Jemma C. Peberdy,[a] Isabelle Meistermann,[a] Laura J. Childs,[a] Christian J. Isaac,[a] Christopher R. Pearmund,[a] Veronika Reudegger,[a] Syma Khalid,[a] Nathaniel W. Alcock,[a] Michael J. Hannon*[a,b] and Alison Rodger*[a]

Received 27th September 2006, Accepted 7th November 2006
First published as an Advance Article on the web 20th November 2006
DOI: 10.1039/b614093a

The enantiomeric resolution of an extended range of di-metallo supramolecular triple-helical molecules are reported. The ligands for all complexes are symmetric with two units containing an aryl group linked *via* an imine bond to a pyridine. Alkyl substituents have been attached in different positions on the ligand backbone. Previous work on the parent compound, whose molecular formula is $[Fe_2(C_{25}H_{20}N_4)_3]Cl_4$, showed that it could be resolved into enantiomerically pure solutions using cellulose and 20 mM aqueous sodium chloride. In this work a range of mobile phases have been investigated to see if the separation and speed of elution could be increased and the amount of NaCl co-eluted with the compounds decreased. Methanol, ethanol and acetonitrile were considered, together with aqueous NaCl : organic mixtures. Effective separation was most often achieved when using 90% acetonitrile : 10% 20 mM NaCl (aq) w/v, which gives scope for scaling up to incorporate the use of HPLC. The overall most efficient (*i.e.* fastest) separation was generally achieved where the cellulose column was packed with 20 mM NaCl (aq) and the column first eluted with 100% acetonitrile, then with 75% ethanol : 25% 20 mM NaCl (aq) until the M enantiomer had fully eluted and finally with 90% acetonitrile : 10% 20 mM NaCl (aq) until the P enantiomer had been collected. The sequence of eluents ensured minimum NaCl accompanying the enantiomers and minimum total solvent being required to elute the enantiomers, especially the second one, from the column. No helicate with a methyl group on the imine bond could be resolved and methyl groups on the pyridine rings also have an adverse effect on resolution.

Introduction

The term helicate was first introduced by Lehn[1] in 1987 to describe the dinuclear $[Cu_2(BP_2)_2]^+$ and trinuclear $[Cu_3(BP_3)_2]^{3+}$ species (BP = oligobipyridine ligands), which were described as polymetallic helical double stranded complexes (see refs. 1–5 and references therein). The first planned strategy for the generation of dinuclear double-stranded helicates may be attributed to Lehn, Sauvage and Ziessel, who prepared the conformationally restricted and pre-organised tetramethyl 2,2′:6′,2″:6″,2‴ quaterpyridine ligand: Me_4(qtpy) and reacted it with copper(I) to produce a dinuclear complex $[Cu_2(Me_4qtpy)_2]^{2+}$with a crystal structure that confirms the helical arrangement of the two strands around the metal ions.[6] Subsequently, it was shown that the helical structure resulted purely from the mismatch between the planar tetradentate donor set of the pyridines together with the pseudo-tetrahedral coordination preference of the metal ion.[7] Although the first triple-stranded helicate was synthesised in 1958 by Stratton and Busch,[8] (the triple-stranded nature was confirmed crystallographically in ref. 9) the first to be structurally characterised was obtained in 1991 by Williams and co-workers who mixed three equivalents of the rigid bis[1-methyl-2-(6′-methyl-2′-pyridyl)benzimidazol-5-yl]methane with two equivalents of cobalt(II) perchlorate.[10] As the ligand cannot wrap around one metal atom, it acts as a bis-bidentate ligand, which, when coordinated to the octahedral Co(II), results in a triple helicate.

One of the aims of our overall research programme has been to produce inexpensive helicate molecules that would bind to the DNA major groove. To date we have shown that both the surface of the ligands and the handedness of the helicate affect their DNA-binding.[11–13] The dinuclear triple-stranded helicate, $[Fe_2(CH_2 \sim H)_3]^{4+}$ (Fig. 1, where the CH_2 denotes the ligand bridging group and ~ H denotes the substituent on the imine), which is the original compound in the series of metal triple helicates discussed

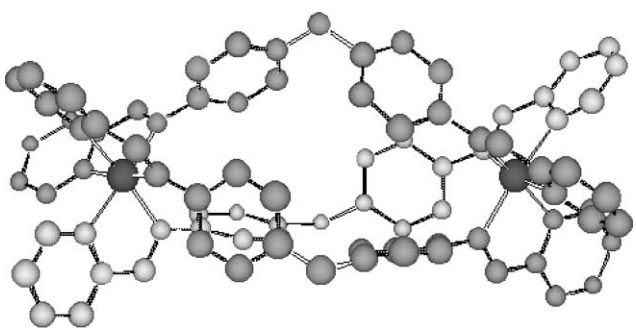


Fig. 1 $[Fe_2(C_{25}H_{20}N_4)_3]Cl_4$, denoted $[Fe_2(CH_2 \sim H)_3]^{4+}$.[15]

[a]*Department of Chemistry, University of Warwick, Coventry, UK CV4 7AL. E-mail: a.rodger@warwick.ac.uk; Tel: +44 24 76574696*
[b]*School of Chemistry, University of Birmingham, Edgbaston, Birmingham, UK B15 2TT, m.j.hannon@bham.ac.uk; Tel: +44 121 414 2527*

† Electronic supplementary information (ESI) available: Characterization data. See DOI: 10.1039/b614093a

in this paper, was obtained as described by Hannon and co-workers using imine ligands.[14–16] Circular dichroism (CD, the difference in absorption of left and right circularly polarised light),[17] linear dichroism (LD, the difference in absorption of light polarised parallel and perpendicular to an orientation axis),[17,18] atomic force microscopy and NMR studies have confirmed that the parent synthetic tetracationic units (Fig. 1) recognise the major groove of DNA and induce remarkable structural effects.[11,12] The complexes are chiral and their helicity has been found to have a significant effect on the DNA binding of the enantiomers.[12,13,19] We have found that the enantiomers of the parent compound $[Fe_2(CH_2 \sim H)_3]^{4+}$ (Fig. 1) can be separated using cellulose (on paper or packed in a column) with 0.02 M aqueous sodium chloride as the eluent.[20]

Our preliminary enantiomeric separation work,[20] was undertaken after an exhaustive attempt to find literature methods that worked for $[Fe_2(CH_2 \sim H)_3]^{4+}$. The main techniques traditionally employed to separate enantiomers involve the use of chiral anions either to achieve diastereo-selective crystallisation or as a component of a chiral chromatographic mobile phase.[21,22] Diastereoselective crystallisations of the supramolecular helicates with a range of chiral anions including sodium *d*-antimonyltartrate and (−)*O*,*O*′-dibenzoyl-(L)-tartrate were attempted without success. Chromatographic techniques proved a little more successful, with some resolution being achieved on SP Sephadex C25 with aqueous disodium (−)*O*,*O*′-dibenzoyl-(L-)-tartrate solution as the mobile phase.[23] Changing to other chiral anions (*e.g.* sodium *d*-antimonyltartrate, *S*-(+)-1,1′-binapthyl-2,2′-diyl hydrogenphosphate) or to the other enantiomer of the chiral dibenzoyl tartrate anion did not seem to offer further improvement. Cyclodextrins, as either mobile [24,25] or stationary phase [26–28] components have been used successfully by others, but proved ineffective in this case.

Frustrated with these conventional chromatographic techniques, ct-DNA was itself used to separate the two enantiomers since there seems to be a significant difference in their binding behaviour. First using a dialysis bag, the inside and outside solutions were each found to have an enantiomeric excess, of opposite handedness. This method had a major downside: only one of the enantiomers could be obtained free, the second one inside the bag being DNA-bound. Moreover the helicate was observed also to attach to the dialysis membrane. DNA was then used as a stationary separation phase using chromatographic paper impregnated with DNA following a previously outlined procedure.[20,29] Washing the paper with ethanol afforded a solution containing an enantio-excess of one of the isomers (as evidenced by a strong circular dichroism spectrum). Solutions containing an excess of the other enantiomer could be obtained by extended washing of the residue with ethanol. Alternatively the DNA-impregnated paper could be used for chromatography using 0.16 M aqueous sodium acetate solution (pH 6.90) as eluent. This led to a spot remaining fixed to the baseline and a spot moving up the paper (associated with some streaking). Washing these two spots off the paper with ethanol gave solutions with opposite signed circular dichroism spectra. However, the solutions obtained from the DNA-impregnated paper technique showed instability (as evidenced by the loss of the MLCT absorption signal over a period of hours), not observed with pure solutions of the compound, indicating that the procedure had introduced some reactive impurities resulting from the reaction of the triple helicate with the DNA fixing agents.

Since paper itself is chiral (being essentially a cellulose derived product) it was decided to examine whether the chiral fibres in paper might themselves be capable of affording separation. The compound was spotted onto Whatman 3MM CHR chromatography paper, and paper chromatography conducted with brine (aqueous 20 mM sodium chloride solution) as the eluent. The purple product split into two distinct spots whose CD signals (when adjusted to the same absorbance intensity) gave equal and opposite circular dichroism spectra.[20] The UV-visible absorption and mass spectra of the solutions are identical, confirming that the chromatography is affording two enantiomers and not two otherwise alternate species. The solutions obtained by this method show no evidence of instability. This separation technique was then scaled up to allow preparative separation simply by using commercially available cellulose particles (~20 micron; Aldrich) as the stationary phase in a column. This allows ready resolution of the di-iron triple helical complexes.[20] The same method was shown to work with triple helicates made with different metals and the same ligand.[20]

In this work we report the synthesis and resolution of an extended range of di-metallo helicates with alkyl substituents at different positions on the ligand backbone (Fig. 2), where a particular aim of the work has been to reduce the amount of NaCl co-eluted with the enantiomers.

Results and discussion

Synthesis and characterisation of the helicates

The chelating ligands in Fig. 2 were all synthesised and complexed with Fe(II) to form di-iron triple helicates, which were precipitated from solution as PF_6 salts. The PF_6 salts of the metal complexes are of high purity but, because the corresponding chloride salts are obtained by anion metathesis they tend to be of lower purity, not in terms of their supramolecular cation but because there are traces of other salts left from the metathesis. The chloride salts are required for water solubility; the PF_6 salts are soluble in organic solvents such as acetonitrile and nitromethane. NMR, CHN, MS and IR data confirm the identity of the complexes. The UV-visible absorbance data obtained for the compounds investigated had very similar spectra with slight variations in wavelength maxima for some of the complexes (see Fig. 3 for typical variation in absorbance spectra).

Application of crystallography to confirm the circular dichroism assignment of the handedness of the first-eluted enantiomer

Coupled-oscillator theory of circular dichroism spectroscopy on the ~280 nm in-ligand exciton band was used to identify the handedness of the first-eluting enantiomer of $[Fe_2(CH_2 \sim H)_3]^{4+}$ as left-handed, or Λ.[30] This enantiomer had originally been labelled (−) on the basis of the sign of the long wavelength metal to ligand charge transfer band.[13,20] To confirm the CD analysis, a sample of the first band to be eluted by a 20 mM NaCl mobile phase from a cellulose column (as described in ref. 20) was collected and converted to the PF_6^- salt following a metathesis in methanol using ammonium hexafluorophosphate. Crystals were grown by diffusing benzene into an acetonitrile solution of the metal complex and the resulting structure is shown in Fig. 4. The

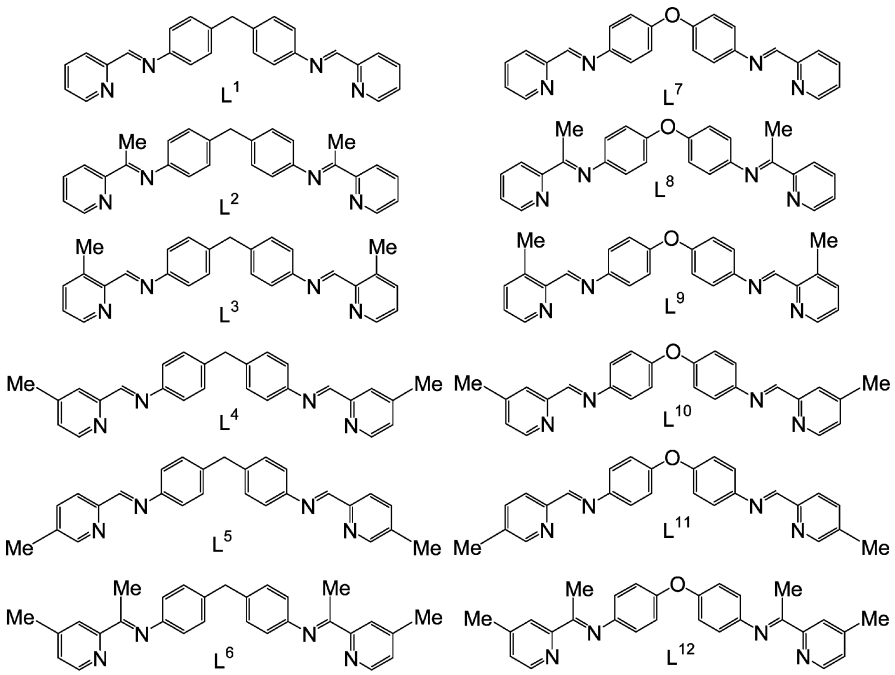


Fig. 2 Schematic diagram of the ligand structures.

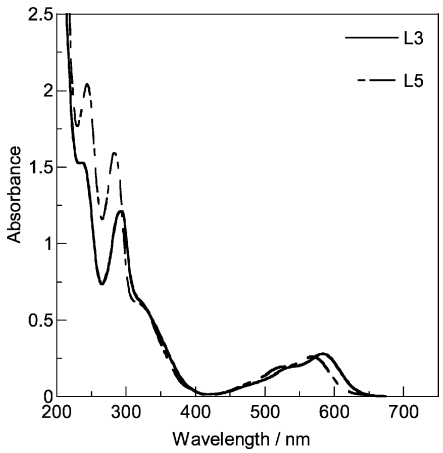


Fig. 3 UV-Visible absorbance spectrum of L_3 and L_5 metal complexes (50 μM, 1 cm path length cuvette) showing typical variations differences in MLCT (~450–650 nm) and in-ligand regions (~250–400 nm) of the spectrum.

first eluting enantiomer, the (−)-enantiomer was indeed found to be the left handed helicate, which is by convention denoted (M). M corresponds directly to Λ absolute configuration at each metal centre. Thus the CD band signs for all complexes can be used to assign the handedness of the compounds.

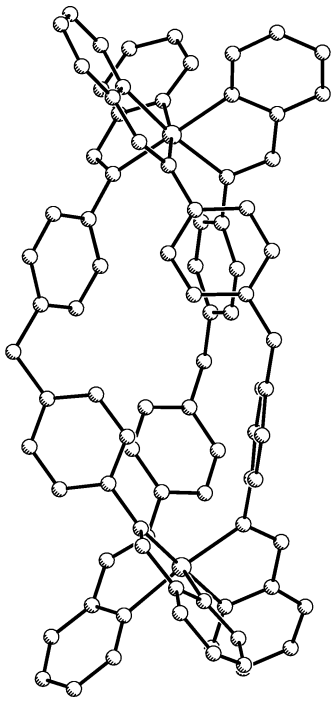


Fig. 4 Crystal structure of the cation of the first eluting (−)-enantiomer of $[Fe_2(L^1)_3]^{4+}$ $4PF_6^-$ showing it to be the M helicate.

Chromatography

Columns were always packed by mixing cellulose with aqueous NaCl (20 M) to form a slurry. The use of any other solvents to pack the column led to significantly poorer resolution so was dismissed in the early stages of the investigation. The chloride salts of the metallo helicates were loaded onto the column from concentrated aqueous solutions and eluted with different mobile phases. The columns resulted in two (if resolution was obtained) purple bands and sometimes a brown–orange one that stayed at the top of the column. Thus the columns simultaneously chemically purified and enantiomerically separated the compounds. Fractions of the purple solutions were collected and the presence of the enantiomers was confirmed using CD (see *e.g.* Fig. 5). The enantiomer giving a negative CD signal at the longest wavelength

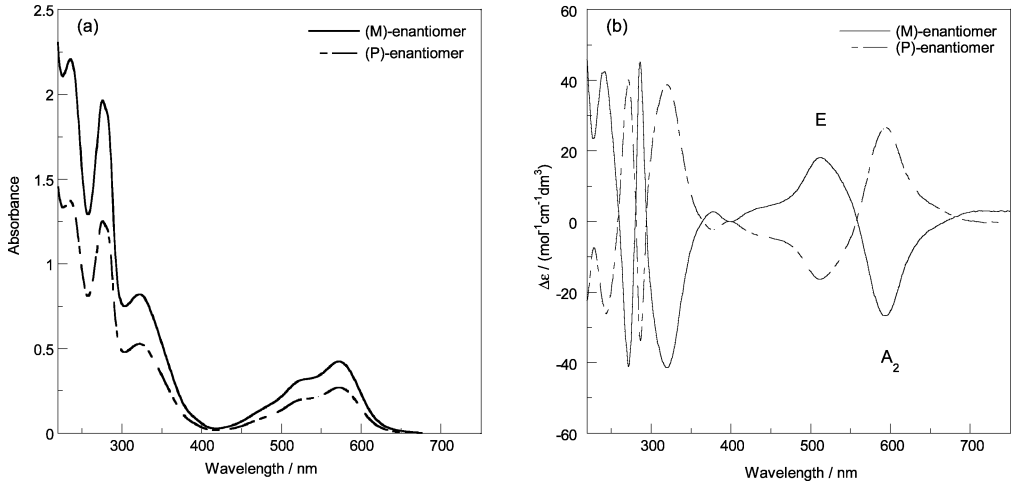


Fig. 5 (a) UV/Visible absorbance spectra (1 cm pathlength) of the two bands eluted from a racemic mixture of $[Fe_2(L^1)_3]^{4+}$ eluted with 90% MeCN : 10% 0.02 M NaCl (aq) w/v. (b) CD spectra of the two solutions were converted to $\Delta\varepsilon$ using $\varepsilon_{574\,nm} = 16900$ mol^{-1} dm^3 cm^{-1} (see methods). For M-$[Fe_2(L^1)_3]^{4+}$ $\Delta\varepsilon_{593.5\,nm} = -26.7$ mol^{-1} dm^3 cm^{-1}.

(592 nm) always eluted first and yielded the most concentrated solutions. As noted above this is the M-helicate.

Aqueous sodium chloride eluent (20 mM) provided a suitable and efficient separation method for the parent compound, $[Fe_2(L^1)_3]^{4+}$ but when applied to the separation of the other metal complexes it gave less satisfactory resolution, with rather dilute solutions (and hence high NaCl content) due to slow elution. A range of different eluents was investigated. 20 mM NaCl proved to be the minimum viable salt concentration for effective separation of the parent compound. Addition of an organic solvent to the aqueous NaCl eluted the components from the column more quickly than the fully aqueous eluent, thus reducing the time taken and producing more concentrated solutions of the enantiomers with less NaCl present. The optimum mobile phase for the parent compound was found to be acetonitrile : aqueous 20 mM NaCl 90 : 10 w/v (Fig. 5). The separation was fast and clearly visible on the column, the enantiomers were eluted in highly concentrated aqueous solution and due to the high organic content, the amount of NaCl in the solution was drastically reduced. Ethanol : aqueous 20 mM NaCl 75 : 25 w/v and methanol : aqueous 20 mM NaCl 80 : 20 were also effective eluents.

The behaviours of the different eluents were such that the most efficient (*i.e.* fastest) separation was generally achieved where the cellulose column was packed with 20 mM NaCl (aq) and the column first eluted with 100% acetonitrile, then with 75% ethanol : 25% 20 mM NaCl (aq) until the M enantiomer had fully eluted and finally with 90% acetonitrile : 10% 20 mM NaCl (aq) until the P enantiomer had been collected. The sequence of eluents ensured minimum NaCl accompanying the enantiomers and minimum total solvent being required to elute the enantiomers, especially the second one, from the column. Acetonitrile provided the fastest movement of the purple bands down the column and reduced the time taken for the enantiomers to be separated compared with the other solvents. This suggests that it decreased the interaction between the enantiomers of the metal complex and the cellulose. The ethanol then facilitates the enantiomeric resolution and the acetonitrile : NaCl mobile phase efficiently removes the P enantiomers from the column. The metal complexes that were effectively separated using this method included complexes formed from L^3, L^5 and L^7 as well as the parent compound $[Fe_2(L^1)_3]^{4+}$ (see Fig. 6 and Table 1 for summary). The P enantiomers of the helicates formed from L^4, L^9, and L^{11} were extremely difficult to remove from the column. Helicates formed from ligands L^2, L^6 and L^{12} showed no separation. Thus it is clear that no universal resolution protocol has been achieved for these helicates. The R_f values from the paper chromatography experiments provide a helpful summary of the separation efficiencies. These and whether enantiomers could be separated are indicated in Table 1.

Discussion and conclusions

Cellulose column chromatography has provided us with a simple and efficient technique for the enantiomeric separation and purification of a range of helicates with different surface topologies. NaCl is required in the eluent to remove the complexes from the column. 20 mM NaCl was also found to be the optimal liquid in which to pack a cellulose column to avoid the cellulose compacting too much. Paper chromatography has also proved a useful method development tool. Although we do not fully understand why cellulose is effective whereas all the literature methods we attempted were ineffective, we have been able to rationalise the effectiveness of the separation of different helicates on elution in terms of the interaction of the helicates with glucose monomers as described in ref. 30. In practice the tertiary structure of the cellulose is certain to play a role too as it is believed to contain helical grooves which could facilitate or hinder the interaction of helicates with the stationary phase.

In this work we have studied the resolution of a sufficiently wide range of compounds that we can make some conclusions about what molecular features switch off the cellulose/helicate enantiomeric discrimination.

(i) All helicates with a methyl group on the imine bond have not been resolved. An imino hydrogen at this position is expected to be acidic and may be able to form CH$\cdots$X hydrogen-bonds to the cellulose structure. Such an interaction is observed in the crystal structure of the complex between the helicate and a 'DNA hexanucleotide'.[31]

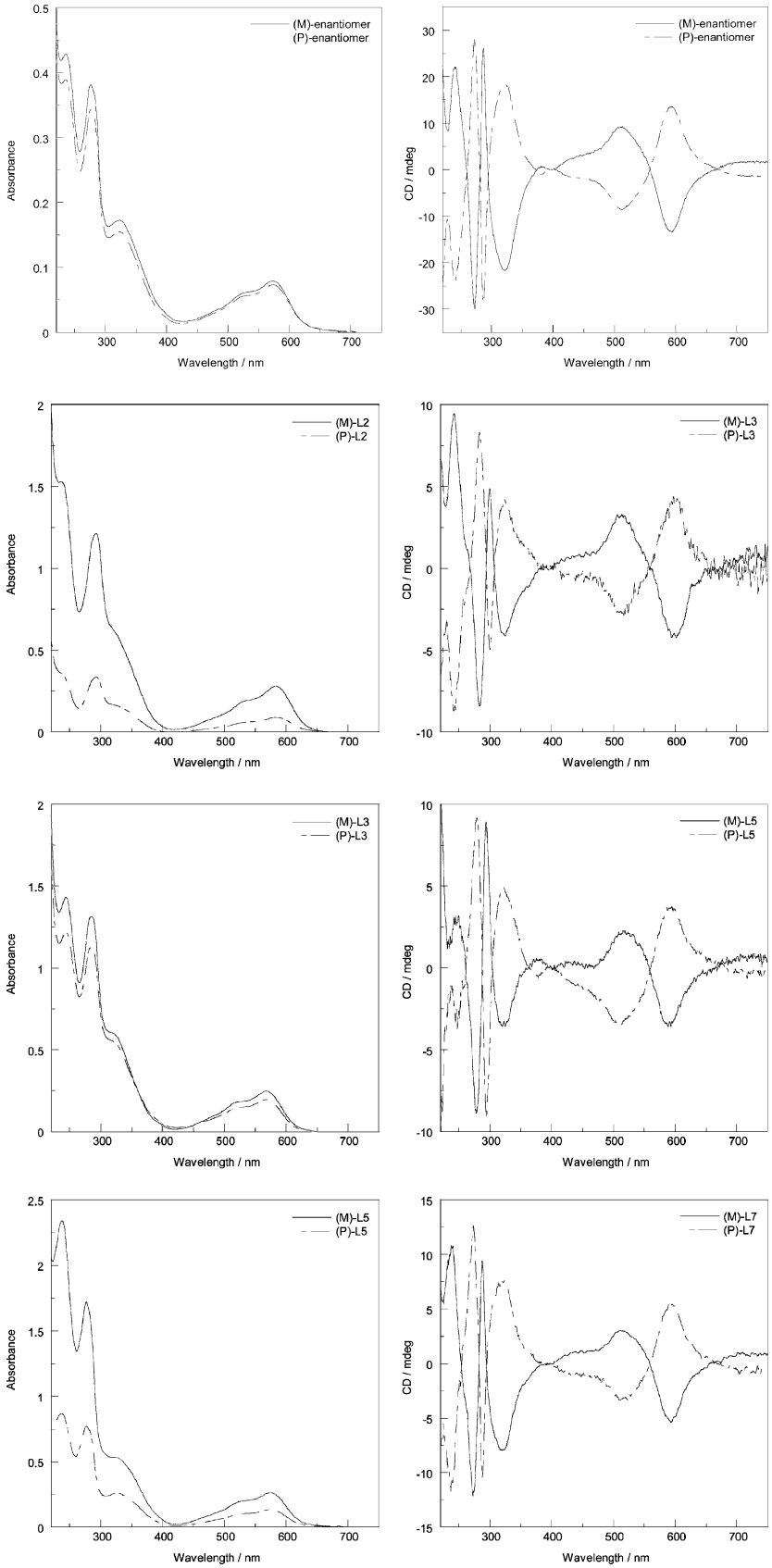


Fig. 6 UV-visible absorbance and CD spectra of the two bands eluted from the column for the di-iron helicates formed from ligands (a) L^1, (b) L^3, (c) L^5 and (d) L^7, eluted using the ideal separating conditions. Spectra were collected in a 1 cm path length cell. A water spectrum was subtracted from each CD spectrum and the spectra were zeroed at 400 nm.

Table 1 Efficiency of column separation (✓✓ denotes good; ✓ denotes sufficient that enantiomerically pure fractions can be collected with care; and ✗ no separation) and paper R_f values (= distance travelled by solvent front/distance travelled by the centre of an analyte band) for the enantiomers of different helicates formed from the ligands indicated in column one. Mobile phases are given in row 1. All mobile phases were based on 20 mM NaCl (aq). The percentage of the organic component is given. MeCN = acetonitrile, MeOH = methanol, EtOH = ethanol

Complex	Separation efficiency/R_f	20 mM NaCl	90% MeCN	80% MeOH	75% EtOH
	✓✓	✓✓	✓✓	✓✓	✓✓
L^1	M	0.60	0.72	0.84	0.88
c	P	0.31	0.56	0.56	0.58
L^2	✗	✗	✗	✗	✗
	✓✓	✓✓	✓✓	✓✓	✓✓
L^3	M	0.63	0.65	0.72	0.84
	P	0.32	0.54	0.53	0.52
	✓	✓	✗	✓	✓
L^4	M	✗	✗	0.49	0.54
	P	✗	✗	0.38	0.39
	✓✓	✓✓	✓✓	✓✓	✓✓
L^5	M	0.64	0.74	0.83	0.97
	P	0.31	0.60	0.59	0.64
L^6	✗	✗	✗	✗	✗
	✓✓	✓	✓	✓	✓
L^7	M	✗	✗	0.94	0.97
	P	✗	✗	0.79	0.83
L^8	✗	✗	✗	✗	✗
	✓	✓	✓	✓	✓
L_9	M	0.73	✗	✗	✗
	P	0.45	✗	✗	✗
	✓	✓	✓	✓	✓
L^{10}	M	0.30	✗	0.50	✗
	P	0.18	✗	0.40	✗
	✓	✓	✗	✓	✓
L^{11}	M	0.48	✗	0.58	✗
	P	0.27	✗	0.33	✗
L^{12}	✗	✗	✗	✗	✗

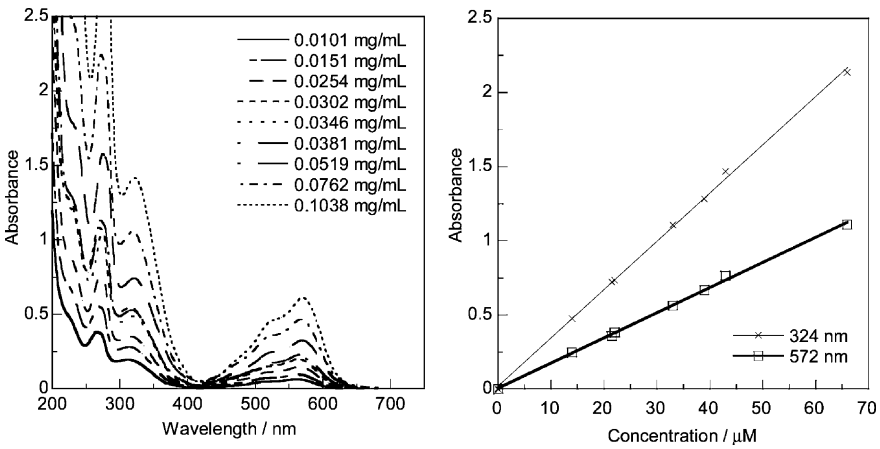


Fig. 7 UV-visible absorbance data for $[Fe(L^1)_3]^{4+}$. Gradient of each line gives the corresponding ε value in the metal–ligand charge transfer and in-ligand regions of the spectrum.

(ii) Methyl groups on the pyridine also have an adverse effect (to varying degrees) in practice, though this may be the result of increased affinity for the column rather than loss of enantiomeric discrimination *per se*. The P-enantiomer of helicates made from L^4, L^9, and L^{11}, for example, proved impossible to elute from the column.

(iii) The experimental conditions we have employed to separate the enantiomers give improved resolution for the helicates with a CH_2 spacer group as opposed to the oxygen spacer group. The modelling results of ref. 30 suggest this is because the CH_2–glucose interaction is more favourable than the O–glucose interaction. Thus O as the spacer group reduces the resolving power of glucose.

Experimental

Materials

The parent compound was synthesised according to previously published methods.[7] All materials were purchased from Sigma-Aldrich unless otherwise stated. 2-Pyridine carboxaldehyde and 2-acetylpyridine (Lancaster) were used without further purification. 3-, 4- and 5-Methyl pyridine carboxaldehyde were prepared from their respective commercially available bromo-methyl pyridines according to a literature procedure.[32] Ultra pure water (18.2 MΩ) was used in all experiments. Cellulose (powder ~ 20 micron) was

used to pack the columns. Celite (type 521) was purchased from Aldrich. 3 Å molecular sieves (crystalline potassium aluminosilicate, 1/16″ pellets) were purchased from BDH Ltd. Methanol, ethanol and acetonitrile were purchased from Fisher Scientific. A range of eluents was prepared by mixing organic solvents with 20 mM NaCl (aq) with organic content ranging from 100% to 50%. The enantiomeric separation was also investigated using less concentrated sodium chloride solutions and with sucrose in the eluent. Sodium antimonyl-L- and D-tartrate were also considered as potential eluent components for the separation of the enantiomers. Sodium antimonyl-L-tartrate (93+%) was purchased from Avocado Research Chemicals. Sodium antimonyl-D-tartrate was synthesised according to literature methods.[33]

Methods

General preparation of the ligands

To a solution of toluene (50 cm^3) containing vacuum dried 3 Å molecular sieves (5 g) and either 4,4′-methylenedianiline or 4,4′-diaminodiphenyl ether (0.010 mol), two equivalents of the relevant pyridine aldehyde/ketone (0.020 mol) were added. The solution was heated under reflux for 24 h. Following filtration through Celite, the solvent was removed *in vacuo* to yield a yellow solid/oil. The ligand was recrystallised from hot ethanol several times to improve purity. The structures of the ligands investigated are summarised in Fig. 2.

Preparation of the iron(II) triple helicates

To a refluxing methanolic solution (50 cm^3) was added either 4,4′-methylenedianiline or 4,4′-diaminodiphenyl ether (0.010 mol) and two equivalents of the relevant pyridine aldehyde/ketone (0.020 mol). This solution was then treated with 2/3 molar equivalents (0.067 mol) of iron(II) chloride tetrahydrate in the minimum amount of methanol. A deep purple coloured solution, characteristic of iron(II) tris-pyridylimine compounds, was formed and reflux was continued for 10 h. The iron(II) triple helicate was precipitated upon the addition of excess ammonium hexafluorophosphate dissolved in methanol. Following filtration, the solid was washed with ether and allowed to dry in a vacuum desiccator. The chloride salt was obtained by anion metathesis in acetonitrile using tetrabutylammonium chloride. The same complexes could be prepared by pre-forming and isolating ligands and then reacting with the metal salt. Characterisation data listings are included in the ESI.†

Parent compound (L^1) triple helicate extinction coefficient determination

An accurate extinction coefficient was determined for $[Fe(L^1)_3]^{4+}$ (L^1=($CH_2 \sim H$)) and these values of ε used to determine the concentrations of all the metal complexes. $\varepsilon_{324\,nm} = 32\,900$ and $\varepsilon_{574\,nm} = 16\,900$ mol^{-1} dm^3 cm^{-1} (Fig. 6). The extinction coefficients determined for $[Fe(L^1)_3]^{4+}$ were assumed to be appropriate for all of the complexes: $\varepsilon_{324\,nm} = 32\,900$ and $\varepsilon_{574\,nm} = 16\,900$ mol^{-1} dm^3 cm^{-1} because it proved extremely challenging to get pure batches of the racemic water soluble chloride salts completely free from salts left from the metathesis. ε values determined for the PF_6^- and Cl^{-1} salts of $[Fe(L^1)_3]^{4+}$ were not the same (the PF_6^- salt requires an organic solvent and is presumably not completely dissociated). Although using the extinction coefficient for L^1 for all compounds is not ideal, at least the resulting error is systematic and data can be easily recalculated if required at a later stage. The difference illustrated in Fig. 3 is only a few percent. The CD spectra of the enantiomers of all the compounds discussed in this work were therefore converted to $\Delta\varepsilon$ assuming all compounds had the same metal to ligand charge transfer extinction coefficient as the parent compound. Solutions were taken to be enantiomerically pure if the CD spectra, expressed as $\Delta\varepsilon$, of the two enantiomers were equal and opposite. It should be noted that any enantiomeric purity calculation is independent of the absolute accuracy of ε.

Table 2 Crystallographic data and structural refinements for (−)-$[Fe_2(L^1)_3]^{4+}$

Empirical formula	$C_{103}H_{89}F_{24}Fe_2N_{14}P_4$
Formula weight	2214.46
Temperature/K	180(2)
Crystal system	Monoclinic
Space group	*C*2
Unit cell dimensions	
a/Å	27.1086(4)
b/Å	10.4113(2)
c/Å	17.4686(3)
α/°	90
β/°	95.292(2)
γ/°	90
Volume/Å^3	4909.19(15)
Z	2
Absorption coefficient	0.463 mm^{-1}
Crystal size	0.8 × 0.075 × 0.075
Reflections collected	16084
Independent reflections	10351 ($R_{int} = 0.0637$)
Data/restraints/parameters	10351/1/665
Goodness-of-fit on F^2	0.890
Final *R* indices [$I > 2\sigma(I)$]	$R_1 = 0.0600$, $wR_2 = 0.1147$
R indices (all data)	$R_1 = 0.1385$, $wR_2 = 0.1361$
Absolute structure parameter	0.008 (19)
Largest diff. peak and hole	0.576 and −0.405 e Å^{-3}

Crystallography

Crystals were successfully grown for the first eluting, (−), enantiomer of $[Fe(L^1)_3]^{4+}$ as the PF_6^- salt by diffusing benzene into an acetonitrile solution of the metal complex. The data are given in Table 2. Crystal data were collected with a Siemens SMART[34] three-circle system with CCD area detector diffractometer. The crystal was held at 180(2) K with the Oxford Cryosystem Cryostream Cooler.[35] Maximum θ was 29.15°; *hkl* ranges −31/35, −13/13, −23/12; 16 084 reflections measured, 10351 unique [R_{int} = 0.0637]. Absorption correction was by semi-empirical from equivalents; maximum and minimum effective transmission factors were: 0.93 and 0.71. There was no crystal decay.

CCDC reference number 622770.

For crystallographic data in CIF or other electronic format see DOI: 10.1039/b614093a

Structure analysis and refinement

Systematic absences indicated either space group *C*2 or *C*2/*m*. The former was chosen on the basis of intensity statistics and shown to be correct by successful refinement. The structure was solved by direct methods using SHELXS [36] with additional light atoms found by Fourier methods, including solvent benzene and

acetonitrile molecules; the complex has crystallographic two-fold symmetry. Hydrogen atoms were added at calculated positions and refined using a riding model with freely rotating methyl groups. Anisotropic displacement parameters were used for all non-H atoms; H-atoms were given isotropic displacement parameters equal to 1.2 (or 1.5 for methyl hydrogen atoms) times the equivalent isotropic displacement parameter of the atom to which the H-atom is attached. The absolute structure was checked by refinement of a delta-f multiplier. Absolute structure parameter $x = 0.008(19)$. Floating origin constraints were generated automatically. An extinction coefficient was refined to 0.00129(16). The weighting scheme was calc $w = 1/[\sigma^2(F_o^2) + (0.0570P)^2]$ where $P = (F_o^2 + 2F_c^2)/3$. Refinement used SHELXTL.[37]

Cellulose paper chromatography

Chromatography cellulose paper (Whatman, 3MM CHR) was cut into 5 cm × 20 cm pieces. The paper was spotted with a concentrated aqueous solution of the metal complex and the enantiomers eluted by capillary action with a 20 mM sodium chloride solution, in which the paper was suspended. Once the solvent front reached nearly to the top of the paper, the paper was removed from the eluent and allowed to dry.

Preparative cellulose column chromatography

Cellulose columns were packed into 2 cm × 30 cm unsintered columns using cellulose particles (~20 micron) as the stationary phase and aqueous 20 mM sodium chloride as the liquid in which the cellulose was suspended for packing. To 6 g of cellulose, 40 cm^3 of 20 mM sodium chloride was added and the solution stirred to a smooth consistency. The column was packed by pouring the aqueous saline suspension of cellulose onto a glass wool pad located just above the stopcock and excess solvent was eluted. The sample, as the chloride salt (the equivalent PF_6 salt is not soluble in aqueous solution), was then loaded onto the column as a saturated aqueous solution (approx. 5 mg mL^{-1}) and the column eluted with the mobile phase of choice. Fraction collection was guided by visual inspection of the profile. If no visible separation of two bands was observed a number of small fractions were collected across the purple band. If separation was apparent, two fractions were collected, one from each purple band.

Analysis cellulose chromatography

A range of mobile phases were investigated to see if the separation and speed of elution could be increased and the amount of NaCl decreased. Methanol, ethanol and acetonitrile were considered, together with aqueous NaCl/organic mixtures. During this method development Pasteur pipettes were adopted as 'mini' columns. This provided an ideal method for investigating packing methods and mobile phases. It reduced the volume of solvent, amount of compound, and time taken for each separation as well as providing disposable columns thus increasing throughput significantly.

Spectroscopy

Circular dichroism (CD) and UV-Visible absorbance spectra were run on Jasco J-715 and Jasco V-550 spectrophotometers, respectively, in 1 cm cuvettes.

Acknowledgements

This work was supported in part by an EU Marie Curie RTN Fellowship (HPRN-CT-2002-00175; JMCAK). M. J. H. is the Royal Society of Chemistry Sir Edward Frankland Fellow 2004–5. We thank the Leverhulme Trust, Socrates-Erasmus (VR from the University of Vienna to Warwick), the EPSRC National Mass Spectrometry Service Centre, Swansea, and the Engineering and Physical Sciences Research Council Life Sciences Initiative for support.

References

1 J.-M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier and D. Moras, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 2565–2569.
2 C. Piguet, G. Bernardinelli and G. Hopfgartner, *Chem. Rev.*, 1997, **97**, 2005–2062.
3 J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCH, Weinheim, 1995.
4 M. Albrecht, *Chem. Rev.*, 2001, **101**, 3457–3497.
5 M. J. Hannon and L. J. Childs, *Supramol. Chem.*, 2004, **16**, 7–22.
6 J.-M. Lehn, J.-P. Sauvage, J. Simon, R. Ziessel, C. Piccinni-Leopardi, G. Germain, J.-P. Declercq and M. Van Meerssche, *Nouv. J. Chim.*, 1983, **7**, 413–420.
7 E. C. Constable, M. J. Hannon, A. Martin, P. R. Raithby and D. A. Tocher, *Polyhedron*, 1992, **11**, 2967–2971.
8 W. J. Stratton and D. H. Busch, *J. Am. Chem. Soc.*, 1958, **80**, 3191–3195.
9 J. Hamblin, A. Jackson, N. W. Alcock and M. J. Hannon, *J. Chem. Soc., Dalton Trans.*, 2002, 1635–1641.
10 A. F. Williams, C. Piguet and G. Bernardinelli, *Angew. Chem., Int. Ed. Engl.*, 1991, **30**, 1490–1492.
11 M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chem., Int. Ed.*, 2001, **40**, 880–884.
12 S. Khalid, M. J. Hannon, A. Rodger and P. M. Rodger, *Chem.–Eur. J.*, 2005.
13 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger and M. J. Hannon, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 5069–5074.
14 M. J. Hannon, C. L. Painting and N. W. Alcock, *Chem. Commun.*, 1999, 2023–2024.
15 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, *Chem. Commun.*, 1997, 1807–1808.
16 A. Rodger, K. J. Sanders, M. J. Hannon, I. Meistermann, A. Parkinson, D. S. Vidler and I. S. Haworth, *Chirality*, 2000, **12**, 221–236.
17 A. Rodger and B. Norden, *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford, 1997.
18 B. Norden, M. Kubista and T. Kurucsev, *Q. Rev. Biophys.*, 1992, **25**, 51–170.
19 S. Khalid, M. J. Hannon, A. Rodger and P. M. Rodger, *J. Mol. Graphics Modelling*, 2006, DOI: 10.1016/j.jmgm.2006.07.004.
20 M. J. Hannon, I. Meistermann, C. J. Isaac, C. Blomme, J. R. Aldrich-Wright and A. Rodger, *Chem. Commun.*, 2001, 1078–1079.
21 A. von Zelewsky, *Stereochemistry of Coordination Compounds*, Wiley, Chichester, 1996.
22 F. R. Keene, *Coord. Chem. Rev.*, 1997, **166**, 121–159.
23 T. J. Rutherford, P. A. Pellegrini, J. R. Aldrich-Wright, P. C. Junk and F. R. Keene, *Eur. J. Inorg. Chem.*, 1998, **11**, 1677–1688.
24 E. Holder, G. Schoetz, V. Schurig and E. Lindner, *Tetrahedron: Asymmetry*, 2001, **12**, 2289–2293.
25 E. Holder, G. Trapp, J. C. Grimm, V. Schurig and E. Lindner, *Tetrahedron: Asymmetry*, 2002, **13**, 2673–2678.
26 D. W. Armstrong, W. DeMond and B. P. Czech, *Anal. Chem.*, 1985, **57**, 481–484.
27 J. M. Green, R. Jones, R. D. Harrison, D. S. Edwards and J. L. Glajch, *J. Chromatogr., A*, 1993, **635**, 203–209.
28 K. Yamanari and M. Nakamichi, *J. Chem. Soc., Chem. Commun.*, 1989, 1723–1724.
29 J. R. Aldrich-Wright, I. Greguric, R. S. Vagg, K. Vickery and P. A. Williams, *J. Chromatogr., A*, 1995, **718**, 436–443.

30 S. Khalid, P. M. Rodger and A. Rodger, *J. Liq. Chromatogr.*, 2005, **28**, 2995–3003.
31 A. Oleksy, A. G. Blanco, R. Boer, I. Usón, J. Aymami, A. Rodger, M. J. Hannon and M. Coll, *Angew. Chem., Int. Ed.*, 2006, **45**, 1227–1231.
32 M. Hatanaka, K. Takahashi, S. Nakamura and T. Mashino, *Bioorg. Med. Chem.*, 2005, **13**, 6763–70.
33 A. Rodger, B. Nordén, P. M. Rodger and P. J. Bates, *Eur. J. Inorg. Chem.*, 2002, **1**, 49–53.
34 *Siemens SMART User's manual*, Siemens Industrial Automation Inc, Madison, WI, USA, 1994.
35 J. Cosier and A. M. Glazer, *J. Appl. Crystallogr.*, 1986, **19**, 105–107.
36 G. M. Sheldrick, *Acta Crystallogr., Sect. A*, 1990, **A46**, 467–473.
37 G. M. Sheldrick, *SHELXTL Ver. 5.1*, Bruker Analytical X-ray Systems, Madison, WI, USA, 1997.

DOI: 10.1002/chem.200601581

Design and Non-Covalent DNA Binding of Platinum(II) Metallacalix[4]arenes

Miguel A. Galindo,[d] David Olea,[c] M. Angustias Romero,[d] Julio Gómez,[c] Pedro del Castillo,[c] Michael J. Hannon,*[a] Alison Rodger,*[b] Félix Zamora,*[c] and Jorge A. R. Navarro*[d]

Dedicated to Professor Bernhard Lippert on the occasion of his 60th birthday

Abstract: A set of cyclic tetranuclear complexes of the metallacalix[4]arene type with formula $[\{Pt(en)(L)\}_4]^{4+}$ (en:ethylenediamine■■ok?■■; **2**: LH = 5-chloro-2-hydroxypyrimidine (5-Cl-Hpymo); **3**: LH = 5-bromo-2-hydroxypyrimidine (5-Br-Hpymo); **4**: LH = 5-iodo-2-hydroxypyrimidine (5-I-Hpymo)) have been obtained from the reaction between *cis*-protected square-planar $[Pt(en)(H_2O)_2]^{2+}$ metal entities and LH in aqueous media. Additionally, the binding properties of **2**, **3**, **4** and their congener $[\{Pt(en)(L)\}_4]^{4+}$ (**1**: LH = 2-hydroxypyrimidine (Hpymo)) with calf thymus-DNA (ct-DNA) have been studied by using different techniques including circular and linear dichroism (CD and LD, respectively) and UV-visible absorbance spectroscopies, gel electrophoresis, fluorescence competitive-binding studies and atomic force microscopy (AFM). The results are consistent with significant non-covalent interactions taking place between the polynuclear cyclic species and ct-DNA. Moreover, gel electrophoresis, linear dichroism titrations and AFM images of ct-DNA with metallacalixarenes show ct-DNA coiling at low metallacalixarene concentrations and upon subsequent increments in metallacalixarene concentration ct-DNA can be seen to uncoil with concomitant formation of long and inflexible ct-DNA structures.■■ok?■■

Keywords: atomic force microscopy · binding studies · DNA · metallacalixarenes · platinum

Introduction

The mechanism of interaction of classical DNA-targeting metal-based drugs generally involves covalent binding to nucleobase moieties and a low degree of selectivity.[1] Therefore, there is a considerable interest in the design of new DNA-binding metal-based drugs that exhibit enhanced selectivity and novel interaction modes, for instance non-covalent interactions that mimic the interaction mode of proteins.[2,3] Recently, a hexacationic ammonium-functionalised bis-calixarene has been observed to bind DNA and has been proposed to do so through binding in the major groove.[4] Metallacalix[*n*]arenes are a class of container molecules that are structurally and functionally related to classical organic calixarenes in which the methylene and phenol rings have been respectively replaced by a metal fragment and a bent nitrogen heterocycle. Like their organic analogues, these systems are able to give a rich variety of host–guest interactions, which include metal-ion, anion and ion-pair recognition.[5] In this regard, we and others have shown that their cationic nature leads to a concomitant high affinity for

[a] Prof. M. J. Hannon
School of Chemistry, The University of Birmingham
Edgbaston, Birmingham, B15 2TT (UK)
Fax: (+44) 1214-147-871■OK?■
E-mail: m.j.hannon@bham.ac.uk

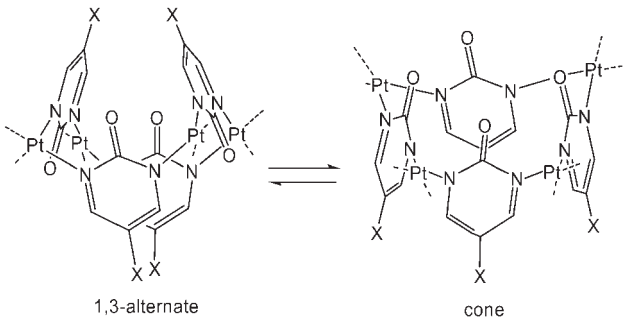
[b] Prof. A. Rodger
Department of Chemistry, University of Warwick
Gibbet Hill Road, Coventry, CV4 7AL (UK)
Fax: (+44) 1203-524-112■OK?■
E-mail: A.Rodger@warwick.ac.uk

[c] Dr. D. Olea, Dr. J. Gómez, Dr. P. del Castillo, Dr. F. Zamora
Facultad de Ciencias, Universidad Autónoma de Madrid
28049 Madrid (Spain)
Fax: (+34) 91-497-4833■OK?■
E-mail: felix.zamora@uam.es

[d] Dr. M. A. Galindo, Dr. M. A. Romero, Dr. J. A. R. Navarro
Departamento de Química Inorgánica Universidad de Granada
Av. Fuentenueva S/N, 18071 Granada (Spain)
Fax: (+34) 95-824-8526
E-mail: jarn@ugr.es

Supporting information for this article is available on the WWW under http://www.chemistry.org or from the author.

anionic substrates.[5–7] In particular, we have observed that metallacalixarenes are especially suited for supramolecular interactions with mononucleotides[8] with some examples of enantioselective recognition processes.[9] These previous results have prompted us to investigate the possible use of metallacalixarenes as selective DNA binding agents. Thus, in this contribution, we describe the non-covalent DNA binding properties of a series of platinum(II)-based metallacalix[4]arenes (Scheme 1).



Scheme 1. Metallacalix[4]arenes $[\{Pt(en)(5\text{-X-pymo})\}_4]^{4+}$ (en:ethylenediamine; **1**: X=H; **2**: X=Cl, **3**: X=Br, **4**: X=I) in their 1,3-alternate and cone conformations. Intermediate conformers are also possible.

Results and Discussion

The novel metallacalix[4]arenes $[\{Pt(en)(5\text{-X-pymo})\}_4](NO_3)_4$ (**2**, **3**, and **4**; see Scheme 1) have been obtained by the reaction of an aqueous solution of $[Pt(en)(H_2O)_2](NO_3)_2$ with the corresponding 5-halo-2-hydroxypyrimidine (5-X-pymoH) derivative to afford, in a single step, the thermodynamically favoured tetranuclear metallamacrocycles. From analysis of the ^{1}H NMR spectroscopy data, it can be deduced that the reaction is quantitative and diagnostic of the N1,N3 bridging coordination mode of 5-X-pymo derivatives, with retention of the original equivalence of H4 and H4′ protons in the 5-X-pymo moieties in a similar way to $[\{Pt(en)(pymo)\}_4](NO_3)_4$ (**1**), which has previously been fully characterised by using single-crystal X-ray crystallography.[10] The most remarkable feature of the ^{1}H NMR spectra is that the H4,H4′ resonances are shifted downfield on passing from the chloro-substituted pymo derivative (δ = 8.43 ppm), to the bromo (δ = 8.49 ppm) and iodo (δ = 8.53 ppm) derivatives, as expected. The ^{1}H NMR spectra also show that the compounds do not decompose in aqueous solution even during long storage periods and over a wide range of pH (1–13), which is in agreement with the highly inert nature of the Pt–N bonds. Variable-temperature ^{1}H NMR experiments (22–85 °C) have been performed for species **1**, **2**, **3** and **4** to explore their conformational flexibility. The results are consistent with conformational flexibility for species **1** and **2** with a unique set of signals observed throughout the whole temperature range. By contrast, above 50 °C, species **3** and **4** show new sets of downfield shifted signals, which agree with a constricted conformational flexibility between the plausible conformers (i.e., 1,2-alternate, 1,3-alternate, partial cone and cone)[10] with predominance of the 1,3-alternate conformer at low temperatures and a mixture of the 1,3-alternate, cone and other intermediate conformers at higher temperatures (Figure 1, Scheme 1). The

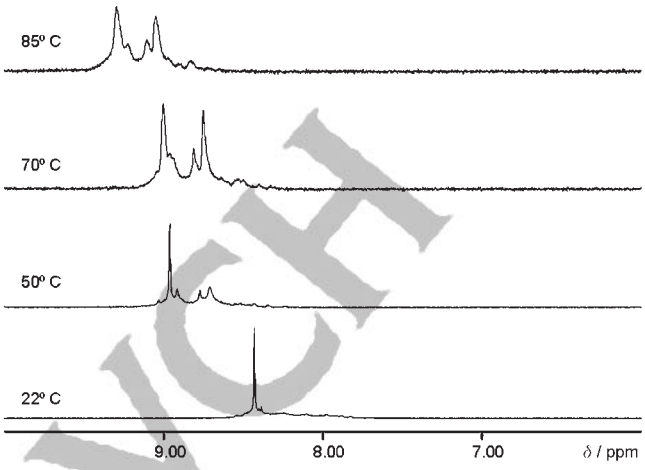


Figure 1. Variable-temperature ^{1}H NMR spectra of **3** in the aromatic region. At high temperature the low- and high-field signals are attributed to the 1,3-alternate and the cone conformer, respectively. The intermediate signals are related to the other conformers (i.e., 1,2-alternate and partial cone).

different behaviour of **1** and **2** relative to **3** and **4** can be attributed to the bulk and heavy nature of the Br and I substituents. The electronic spectra of **1**–**4** each show a broad ligand-to-ligand absorption band in the 300–390 nm region, centred at 323 nm for **1** (ε = 14400 $mol^{-1}dm^3cm^{-1}$), 348 nm for **2** (ε = 34000 $mol^{-1}dm^3cm^{-1}$), 350 nm for **3** (ε = 27500 $mol^{-1}dm^3cm^{-1}$), and 358 nm for **4** (ε = 13000 $mol^{-1}dm^3cm^{-1}$).

DNA binding studies: UV-visible absorbance, circular dichroism (CD) and flow linear dichroism (LD) titration experiments between the compounds and calf thymus DNA (ct-DNA) have been carried out in aqueous solution with NaCl (20 mM) and sodium cacodylate buffer (1 mM) while keeping the ct-DNA concentration constant (300 μM).

The UV-visible titration experiments of ct-DNA with **1** (see the Supporting Information) show no significant changes in the 300–390 nm ligand-centred complex bands on addition of the ct-DNA. This is consistent with the structure of the metallacalixarene remaining unchanged and thus implies that if there is an interaction of the tetracationic species **1** with ct-DNA it should be of supramolecular nature rather than covalent DNA platination. This is consistent with the chemical inertness of the Pt–N bonds in the **1**–**4** series, which prevent any covalent binding to ct-DNA.[11] The region in which the ct-DNA absorbs changes in intensity as the titration experiment progresses. The variation in the intensity of this ct-DNA band can be correlated with the changes in base–base stacking. Thus, at low concentrations

of metallacalixarene there is a decrease in the intensity of the band associated with the ct-DNA, which appears to exclude intercalation as a binding mode at low metallacalixarene loading. Intercalation reduces the flexibility of the ct-DNA and is usually associated with an increase in the 260 nm absorption band, but at low concentrations of metallacalixarene a certain degree of coiling can be induced by supramolecular interactions.■■ok?■■

To gain further information, we also recorded CD and LD spectra. The complexes **1** to **4** have no intrinsic CD signals as they are achiral. Any CD signals that arise in the spectroscopic regions of the complex are therefore a consequence of their interaction with the ct-DNA. Linear dichroism (LD) is the difference in absorption of linearly polarised light both parallel and perpendicular to a chosen plane and can be used to probe the orientation of molecules. Long molecules, such as DNA (minimum length of ≈250 base pairs) can, in a flow Couette cell, be orientated through viscous drag.[2] The linearly polarised light is incident radial to the flow cell and perpendicular to the flow direction. Small unbound molecules are not orientated in the experiment and show no signal. Similarly molecules bound randomly to the ct-DNA show no signal. However, molecules bound in a specific orientation with respect to the ct-DNA will show a signal. Complexes **1**–**4** are too small to be orientated and thus show no intrinsic signal. Any signals that arise in the spectroscopic regions of the complex after the addition of ct-DNA, therefore, indicate binding of the complex to the ct-DNA in a specific orientation(s).

For each of the metallacalix[4]arenes **1**–**4** we observed bands in the 300–390 nm region in both the LD (Figure 2) and CD spectra (see the Supporting Information). These results indicate that each of these complexes binds to ct-DNA and does so in a specific orientation(s), not randomly.[2] Moreover, the band that arises in the 350–370 nm region suggests that the pyrimidine heterocycles might be orientated less than 54.7° from the helix axis. This often means that these systems are slotted in a groove, which from size considerations should be the ct-DNA major groove.[4] The ct-DNA LD bands (220–300 nm) confirm that the ct-DNA remains in the B-DNA conformation, however, some structural changes in ct-DNA are suggested by the decrease in the ct-DNA LD band at 260 nm upon metallacalix[4]arene addition. The ct-DNA LD signal is consistent with a non-intercalative mode of interaction (intercalation reduces the flexibility of the ct-DNA and is usually associated with an increase in the 260 nm LD band). Indeed at low loading there is very little change in the ct-DNA LD signal. This contrasts with the dramatic loss of ct-DNA LD signal seen when tetracationic supramolecular cylinders bind and coil ct-DNA even at very low loading.[2] The results are similar along the **1**–**4** series, however, these conclusions are hampered for compounds **3** and **4** owing to some degree of precipitation.

Fluorescence competitive-binding assays have also been performed with ethidium bromide (EB) to evaluate the strength of the interaction of the metallacalixarenes **1** to **4** with ct-DNA (Figure 3). It should be noted that in contrast

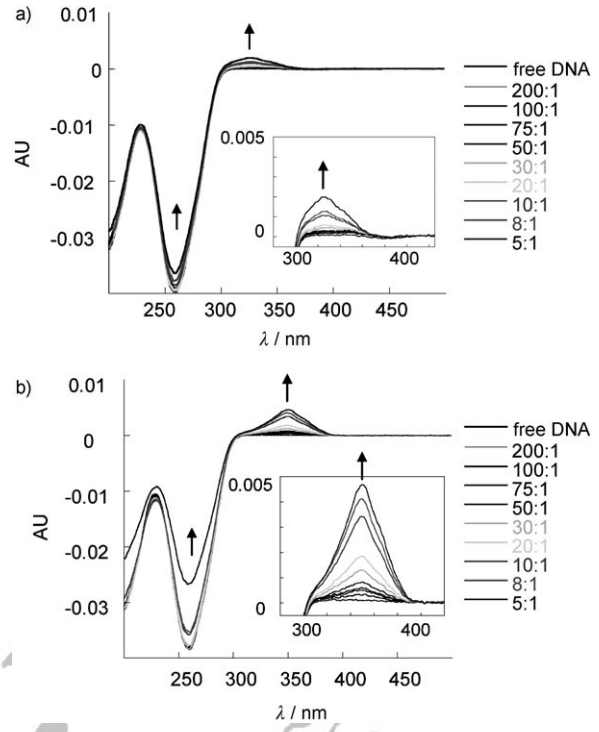


Figure 2. LD spectra of free ct-DNA (300 μM) and in the presence of **1** (a) and **2** (b). ct-DNA/metallacalixarene mixing ratios range from 200:1 to 1:1. Inset enlargements ×4.

to previous experiments, precipitation does not take place in these experiments (where a lower ct-DNA concentration is used). Although, the UV-visible absorbance, LD and CD spectroscopic studies revealed that complexes **1** and **2** interact with ct-DNA, the strength of this interaction is not large enough to displace intercalated EB (Figure 3a), which is indicative of a lower association constant (K_{ass}) than that of EB. Complexes **3** and **4**, however, interact strongly with ct-DNA. This is manifested by the quenching of the band emission centred at 600 nm, which is associated with ct-DNA intercalated with EB, and is a consequence of the displacement of EB by the metallacalixarene (Figure 3b), which should be indicative of a association K_{ass} value in the same range as that of EB. It should also be noted that emission decreases upon complex addition until a metallacalixarene concentration of ≈2 μM (ratio ct-DNA base pairs/metallacalixarene, 3:1) is reached, indicating the saturation of ct-DNA binding sites and a high K_{ass} for **3** and **4** (Figure 3b) and in the Supporting Information). The different behaviour of **3** and **4** relative to **1** and **2** might be related to their different conformational behaviour (see NMR data discussed above). Thus, we presume that the interaction of the metallacalixarene with ct-DNA is enhanced by the fixation of a particular conformer of those possible for **3** and **4**, driven by the optimal fitting to the ct-DNA surface ■■ok?■■. The higher binding strength of **3** and **4** may also be the reason for their lower precipitation concentration from the absorbance, CD and LD experiments.

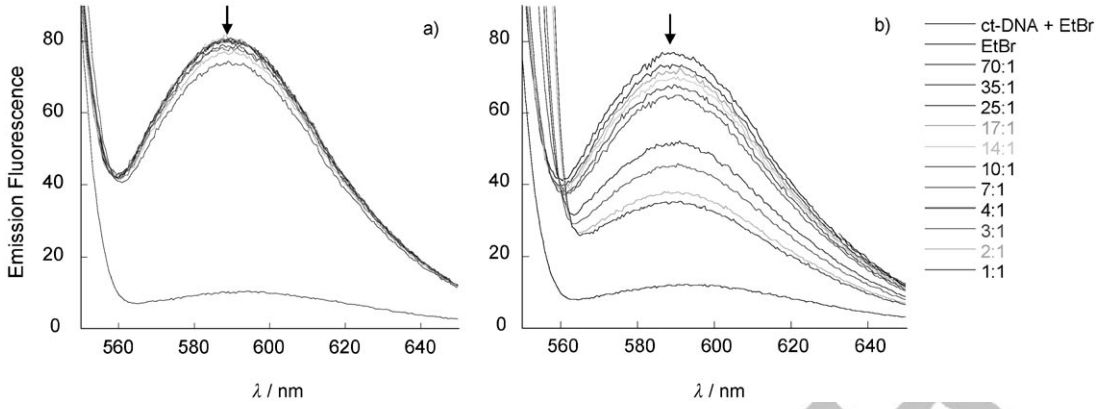


Figure 3. Fluorescence quenching of ethidium (5 μM) bromide and ct-DNA (4 μM) in the presence of **2** (a) and **4** (b). ct-DNA base/metallacalixarene mixing ratios range from 70:1 to 1:1.

Gel electrophoresis experiments were performed on the ct-DNA/**1** system to follow the ct-DNA migration when it is bound to the complexes. The experiments show a smooth decrease in the ct-DNA migration rate for the ct-DNA/**1** ratios from 200:1 to 30:1 (see the Supporting Information). At higher complex concentrations the migration rate diminishes dramatically with no migration at all at ratios lower than 5:1. The smooth decrease in the ct-DNA migration rate can be attributed to the reduction of its negative charge as a consequence of the interaction with the tetracationic metallacalixarenes. However, the dramatic decrease of the ct-DNA migration rate at higher complex concentrations could be related to the formation of inflexible, long ct-DNA structures (see the AFM experiments below).

AFM studies: To clarify the concentration dependent ct-DNA conformational changes induced by the metallacalixarenes we have carried out atomic force microscopy (AFM) experiments at different ct-DNA/metallacalixarene mixing ratios.

An image of several molecules of calf thymus DNA adsorbed on a mica surface by using $MgCl_2$ is shown in Figure 4 (blank sample, see the Experimental Section for details). The molecule height is about 1 nm and has a persistence length of ≈50 nm, these dimensions are typical values for AFM measurements.[12,13] The length of the ct-DNA is not well defined, ranging from 300 to 2000 nm.

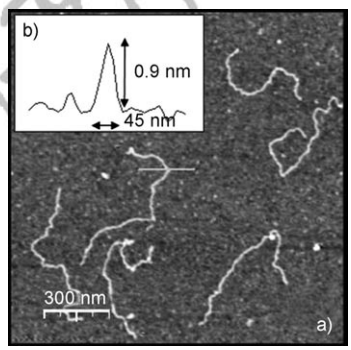


Figure 4. AFM topographic image showing ct-DNA double strands adsorbed on a mica substrate by using the classical method of adding $MgCl_2$ to a buffer solution with ct-DNA.

When ct-DNA was incubated with **1** at a mixing ratio of 30:1, we observed that on the mica surface and along the ct-DNA molecules, spots had appeared (Figure 5). These spots

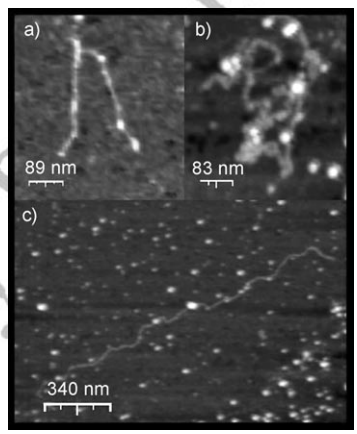


Figure 5. AFM topographic images of ct-DNA incubated with **1** (30:1 mixing ratio) adsorbed on mica. The white dots along the ct-DNA are attributed to ct-DNA/metallacalixarene supramolecular assemblies with a concomitant increase of the height of the strand from 0.6 nm to ≈2 nm.

have been attributed to the metallacalixarenes. In this case (and in contrast to the 10:1 mixing ratio, see below) a certain degree of bending or coiling can be seen. In addition, the AFM topographic images show a marked drop of the persistence length with respect to free ct-DNA. This drop in the persistence length is also in agreement with some small amounts of bending or coiling, although no dramatic kinks or coils are observed; consistent with the LD results. Reductions in ct-DNA persistence length by multivalent cations are not unusual and have been attributed to the bending of the ct-DNA.[14] Finally, it should be noted that the height of the ct-DNA strands is in the typical 0.6–0.8 nm range, which

increases to 2 nm in the regions where the spots attributed to metallacalixarenes bound to the ct-DNA are observed.

In contrast to the lower loadings, if ct-DNA is incubated with larger loadings of metallacalixarenes (ct-DNA/**1** in a 10:1 ratio) the AFM topographic images show inflexible, long ct-DNA filaments with lengths often higher than 15 μm (Figure 6). Its persistence length is, on average, much longer

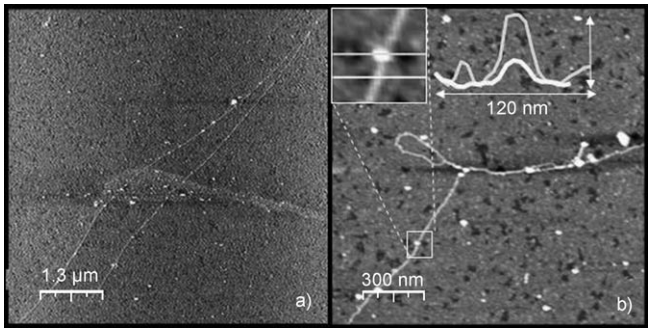


Figure 6. AFM topographic images of ct-DNA incubated with **1** (10:1 mixing ratio) and adsorbed on mica. Long structures with high persistence length are observed. a) ≈15 μm ct-DNA strands. b) ct-DNA strand with a typical height of 0.8 nm. The white dots along the ct-DNA are attributed to ct-DNA/metallacalixarene supramolecular assemblies.

than in blank ct-DNA. Strands of ct-DNA from these larger loadings had a height of 0.8 nm, which is not significantly different from the ct-DNA height observed in the blank sample seen in Figure 6b. The inset in Figure 6b shows a 2.5 nm high spot in the centre of a straight ct-DNA fibre. We attribute this spot to a metallacalixarene bound to the ct-DNA. In addition, similar spots are shown to be unspecifically adsorbed on a mica surface in Figure 6b.

The structural changes that occur when ct-DNA is incubated with **4** (ct-DNA/**4** in a 10:1 mixing ratio) are similar but more dramatic than the changes seen for ct-DNA/**1** (Figure 7 and Figure S3 in the Supporting Information). This is consistent with **4** having a higher affinity for ct-DNA than

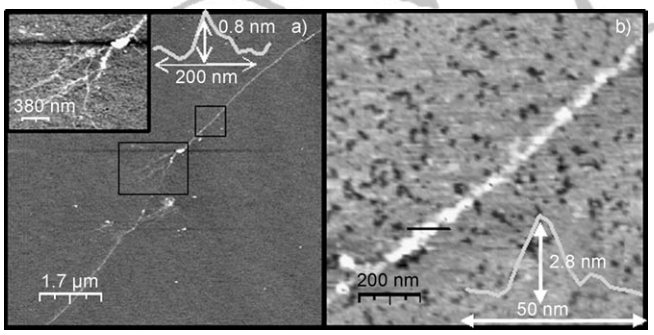


Figure 7. a) AFM topographic image of ct-DNA incubated with **4**. The inset shows several ct-DNA strands with a height of 0.8 nm indicating that the longer structure is a bundle with a height of 2.6 nm, much higher that the typical height measured with AFM for individual ct-DNA strands. b) A magnified image taken from the smaller square marked in (a).

1, which matches the conclusions drawn from the spectroscopic studies. Very large structures can now be seen on the mica surface. These structures are likely formed by a mixture of ct-DNA double strands and metallacalixarenes that result in long fibres that in some cases (Figure 7b inset) give rise to large two-dimensional networks.

Shown in Figure 7b is an enlargement of the marked region of Figure 7a. In this image a quasi one-dimensional granular structure is clearly seen. The image suggests that ct-DNA/metallacalixarene assemblies give rise to the formation of long filaments of high persistence length. It is remarkable that in the central region of Figure 7a several strands of what are presumably individual ct-DNA molecules can be seen coming out from the fibre (see inset of Figure 7a).

Conclusions

We have observed clear non-covalent interactions between metallacalix[4]arenes and calf thymus DNA (ct-DNA). The non-covalent interactions were expected, based on the chemically inert nature of the platinum-containing metallacalix[4]arenes, and were deduced from the spectroscopic and AFM studies. The AFM images clearly show that the assayed compounds **1** and **4** interact with ct-DNA to produce supramolecular ct-DNA/metallacalixarene assemblies that induce conformational changes in the ct-DNA. This type of supramolecular interaction does not actively reduce the flexibility of the ct-DNA at low loadings (perhaps causing some small amounts of kinking or coiling), but this situation is reversed at higher loading with the formation of inflexible ct-DNA filaments. We presume that low concentrations of tetracationic metallacalixarene start to quench the ct-DNA charge and may enhance flexibility. This situation is reversed at higher complex concentrations, probably as a consequence of the occupation of a large number of ct-DNA binding sites, with a concomitant ct-DNA uncoiling, reduction in flexibility and subsequent aggregation into higher-order structures.

The contrast with previous studies on the effects of tetracationic cylinders on ct-DNA is revealing. Both classes of agents are tetracations that have extensive π surfaces and both bind non-covalently to ct-DNA. The cylinders induce dramatic intramolecular ct-DNA coiling at both low and high loadings, whereas these metallacalixarene agents do not induce dramatic coiling effects but rather reduce the flexibility of and aggregate the ct-DNA. Tetracationic spermines condense, and may aggregate, ct-DNA but do not impart a reduction in flexibility or lead to fibre formation.

This not only highlights the importance of the precise cylindrical structure of the previously reported supramolecular ct-DNA-coiling agents, but implies that different supramolecular drug designs might be used to induce different DNA structural effects. It remains to be seen whether there could be any biological relevance or application of the DNA conformational changes induced by this kind of supramolec-

ular interaction described herein and this is the subject of planned future studies.

Experimental Section

Materials: 2-pyrimidinol·HCl (2-Hpymo·HCl) was purchased from Aldrich. 5-chloro-2-hydroxypyrimidine,[15] 5-bromo-2-hydroxypyrimidine,[16] 5-iodo-2-hydroxypyrimidine,[16] and [{Pt(en)(pymo)}$_4$] $(NO_3)_4$ (**1**)[10] were prepared according to literature methods.

Preparation of [{Pt(en)(L)}$_4$](NO$_3$)$_4$ (2: LH=5-Chloro-2-pyrimidinolate; 3: LH=5-Bromo-2-pyrimidinolate; 4: LH=5-Iodo-2-pyrimidinolate): [$PtCl_2$(en)] (1 mmol) was suspended in an aqueous solution of $AgNO_3$ (2 mmol, 25 mL). The suspension was stirred in the dark at 60 °C overnight. The resulting mixture was kept at 4 °C for some minutes before the AgCl precipitate was filtered off. The resulting filtrate was mixed with an aqueous solution of the corresponding HL (1 mmol in 30 mL), the pH was adjusted to 5.5 by means of NaOH (1 M) and the mixture was allowed to react at 60 °C for 5 h. The mixture was reduced to 10 mL by means of a rotary evaporator to afford the corresponding metallacalixarene.

Complex 2: Yield: 80 %; ^{1}H NMR (400 MHz, D_2O, 25 °C, TMA■■what does TMA stand for?■■): δ = 2.5–2.7 (m, 16 H; en), 8.43 ppm (s, 8 H; H4,H4′-5Cl-pymo); elemental analysis calcd (%) for $C_{24}N_{20}H_{40}O_{16}Pt_4Cl_4\cdot(H_2O)_8$ (1930.96 g mol^{-1}): C 14.92, H 2.92, N 14.50; found: C 14.85, H 2.40, N 14.54.

Complex 3: Yield: 80 %; ^{1}H NMR (400 MHz, D_2O, 25 °C, TMA■■): δ = 2.5–2.7 (m, 16 H; en), 8.49 ppm (s, 8 H; H4,H4′-5Br-pymo); elemental analysis calcd (%) for $C_{24}N_{20}H_{40}O_{16}Pt_4Br_4\cdot(H_2O)_8$ (2108.76 g mol^{-1}): C 13.67, H 2.67, N 13.28; found: C 13.50, H 2.39, N 13.15.

Complex 4: Yield: 75 %; ^{1}H NMR (400 MHz, D_2O, 25 °C, TMA■■): δ = 2.5–2.7 (m, 16 H; en), 8.53 ppm (s, 8 H; H4,H4′-5I-pymo); elemental analysis calcd (%) for $C_{24}N_{20}H_{40}O_{16}Pt_4I_4\cdot(H_2O)_8$ (2296.77 g mol^{-1}): C 12.55, H 2.45, N 12.19; found: C 12.59, H 2.25, N 12.34.

DNA binding studies: Calf-thymus DNA (ct-DNA) was purchased from Sigma/Aldrich. The ct-DNA was dissolved in water without any further purification and kept frozen until the day of the experiment. The ct-DNA concentration (moles of bases per litre) was determined spectroscopically by using the molar extinction coefficients at the maximum of the long-wavelength absorbance (ct-DNA ε_{258} = 6600 cm^{-1} mol^{-1} dm^3). Concentrations of stock solutions of the metallacalix[4]arenes were determined from accurately weighed samples of this material. A stock sodium cacodylate buffer (100 mM) was prepared by mixing a 50 mL solution of sodium cacodylate (0.2 M, 4.24 g of $Na(CH_3)_2AsO_2\cdot 3H_2O$ in 100 mL) with 9.3 mL of hydrochloric acid (0.2 M), and diluting to a total of 100 mL. ■■Stock solutions of the metallacalix[4]arene (500 μM) were prepared and the salt, sodium chloride, and all ct-DNA experiments were conducted in sodium cacodylic buffer (1 mM) and NaCl (20 mM)■■pls clarify this sentence■■. Spectroscopic titration series experiments keeping the ct-DNA concentration constant were undertaken by adding the salt, buffer, water and metallacalix[4]arene to the ct-DNA. The circular dichroism (CD) spectra were produced by using a Jasco J-715 spectropolarimeter. Linear dichroism (LD) experiments were performed in a Jasco J-715 spectropolarimeter adapted for LD using a 1 mm pathlength Couette flow cell.

Ethidium bromide (EB) displacement by the metallacalix[4]arenes was calculated by measuring the quenching of the EB fluorescence as it leaves the protection of the ct-DNA.

A ct-DNA/salts/buffer solution with EB (ct-DNA/EB 4:5, 4 μM:5 μM) was prepared. The emission spectrum was recorded as a function of metallacalix[4]arene concentration by using a Perkin–Elmer Luminescence spectrometer (LS50b) and the metallacalix[4]arene concentration was slowly increased for ct-DNA/metal-complex ratios from 70:1 to 1:1 keeping the ct-DNA and EB concentrations constant. After each addition the fluorescence and UV-visible spectra were recorded (parameters: emission: 600 nm; excitation: 540 nm; excitation slit: 10.0 nm■■; emission slit: 15.0 nm■■).

The gel electrophoresis was done using a HE99X Max submarine electrophoresis tray and EPS 3051 power supply from Amersham, UK. The power used was 7.5 V per cm of gel and run in 1X TAE buffer. The gel was 1 % agarose in 1X TAE buffer. The dye mixture was 0.5 % bromophenol blue and 0.5 % xylene cyanole in 30 % glycerol/70 % water. It was developed by soaking for 30 min in a bath of ethidium bromide (0.05 μg mL^{-1}), followed by 5 min in a magnesium chloride (0.01 M) and irradiated at 312 nm for visualisation.

Atomic force microscopy (AFM):

Adsorption of blank calf thymus DNA: Samples were prepared by depositing a drop (10 μL) of ct-DNA solution (0.001 g L^{-1}) containing $MgCl_2$ (4 mM) onto a recently cleaved muscovite mica sheet (Electron Microscopy Sciences). After adsorption for 2 min at room temperature, the samples were gently rinsed with Milli-Q quality water and dried with nitrogen.

*Incubation of ct-DNA with **1a** and **4***: Solutions of ct-DNA (0.01 g L)$^{-1}$ were incubated at 25 °C with **1a** and **4** (10 basepairs/metallacalixarene) for one hour.

Adsorption of ct-DNA–platinumcalix[4]arenes: The ct-DNA–metallacalixarene solutions were diluted 10 times. A drop of these solutions (10 μL) was deposited onto a recently cleaved sheet of mica for 2 min. The sample was then rinsed and dried as described above.

AFM imaging: AFM images were acquired in dynamic mode using a Nanotec Electronica system (http://www.nanotec.es). Olympus cantilevers were used with a nominal force constant of 0.75 N m^{-1}.

Acknowledgements

The authors acknowledge the Spanish Ministry of Education and Science for grants CTQ2005-00329/BQU, MAT2004-05589-C02-01/02 and NISAN/IST-2001-38052 and for PhD fellowships (M.A.G.). We also thank Junta de Andalucía and COST-D20 Action for funding. This work was conducted within the framework of COST D20 (metal complexes in anti-cancer and anti-viral therapy) and working group D20/0010/02 (non-covalent DNA recognition), and supported an exchange of researchers (M.A.G.).

[1] See, for example: *Cisplatin, Chemistry and Biochemistry of A Leading Anti-Cancer Drug* (Ed.: B. Lippert), Wiley-VCH, Weinheim, **1999**, and references therein.

[2] a) C. Uerpmann, J. Malina, M. Pascu, G. J. Clarkson, V. Moreno, A. Rodger, A. Grandas, M. J. Hannon, *Chem. Eur. J.* **2005**, *11*, 1750–1756; b) M. J. Hannon, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 880–884; c) I. Meistermann, V. Moreno, M. J. Prieto, E. Molderheim, E. Sletten, S. Khalid, P. M. Rodger, J. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074; d) L. J. Childs, J. Malina, B. E. Rolfsnes, M. Pascu, M. J. Prieto, M. J. Broome, P. M. Rodger, E. Sletten, V. Moreno, A. Rodger, M. J. Hannon, *Chem. Eur. J.* **2006**, *12*, 4919–4927; e) S. Khalid, M. J. Hannon, A. Rodger, P. M. Rodger, *Chem. Eur. J.* **2006**, *12*, 3493–3506; f) A. Oleksi, A. G. Blanco, R. Boer, I. Uson, J. Aymami, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem.* **2006**, *118*, 1249–1253; *Angew. Chem. Int. Ed.* **2006**, *45*, 1227–1231; g) A. C. G. Hotze, B. M. Kariuki, M. J. Hannon, *Angew. Chem.* **2006**, *118*, 4957–4960; *Angew. Chem. Int. Ed.* **2006**, *45*, 4839–4842.

[3] a) P. B. Glover, P. R. Ashton, L. J. Childs, A. Rodger, M. Kercher, R. M. Williams, L. De Cola, Z. Pikramenou, *J. Am. Chem. Soc.* **2003**, *125*, 9918–9919; b) L. A. Levine, C. M. Morgan, K. Ohr, M. E. Williams, *J. Am. Chem. Soc.* **2005**, *127*, 16764–16765.

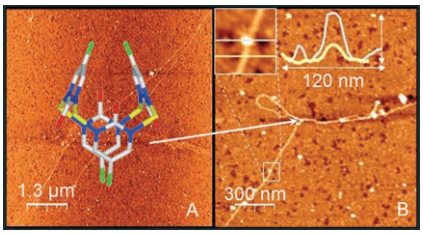
[4] R. Zadmard, T. Schrader, *Angew. Chem.* **2006**, *118*, 2769–2772; *Angew. Chem. Int. Ed.* **2006**, *45*, 2703–2706.
[5] J. A. R. Navarro, E. Barea, M. A. Galindo, J. M. Salas, M. A. Romero, M. Quirós, N. Masciocchi, S. Galli, A. Sironi, B. Lippert, *J. Solid State Chem.* **2005**, *178*, 2436–2451, and references therein.
[6] S. Y. Yu, H. Huang, H. B. Liu, Z. N. Chen, R. B. Zhang, M. Fujita, *Angew. Chem.* **2003**, *115*, 710–714; *Angew. Chem. Int. Ed.* **2003**, *42*, 686–690.
[7] Z. Q. Qin, M. C. Jennings, R. J. Puddephatt, *Inorg. Chem.* **2002**, *41*, 3967–3974;.
[8] M. A. Galindo, S. Galli, J. A. R. Navarro, M. A. Romero, *Dalton Trans.* **2004**, 2780–2785.
[9] E. Barea, J. A. R. Navarro, J. M. Salas, M. Quirós, M. Willermann, B. Lippert, *Chem. Eur. J.* **2003**, *9*, 4414–4421.
[10] J. A. R. Navarro, E. Freisinger, B. Lippert, *Inorg. Chem.* **2000**, *39*, 2301–2305.
[11] Previous studies on the interaction of mononucleotides with palladium-containing metallacalixarenes have shown that the supramolecular recognition is followed by coordination of Pd^{II} to the mononucleotide with the concomitant decomposition of the metallamacrocycle.
[12] H. G. Hansma, R. L. Sinsheimer, M. Q. Li, P. K. Hansma, *Nucleic Acids Res.* **1992**, *20*, 3585–3590.
[13] J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker, N. H. Dekker, *Biophys. J.* **2005**, *88*, 2737–2744.
[14] H. Schutz, K. E. Reinert, *J. Biomol. Str. Dyn.* **1991**, *9*, 315–329.
[15] D. G. Crosby, R. V. Berthold, *J. Org. Chem.* **1960**, *25*, 1916–1919.
[16] S. M. N. Efange, E. M. Alessi, H. C. Shih, Y. C. Cheng, T. J. Bardos, *J. Med. Chem.* **1985**, *28*, 904–910.

Received: November 6, 2006
Published online: ■■ ■■, 2007

DNA Binding

M. A. Galindo, D. Olea,
M. A. Romero, J. Gómez,
*P. del Castillo, M. J. Hannon,**
A. Rodger, F. Zamora,**
*J. A. R. Navarro** ..... ▮▮▮▮–▮▮▮▮

Design and Non-Covalent DNA Binding of Platinum(II) Metallacalix[4]arenes



Charming the DNA snake: Supramolecular binding of tetracationic metallacalixarenes to calf thymus DNA (ct-DNA) induces limited-DNA coiling in a first stage and subsequent uncoiling and reduction in flexibility upon increasing the metallacalixarene complex concentration. These effects have been studied, in part, by atomic force microscopy (see figure).

doi:10.1016/j.jmb.2007.03.025
Publication 120
J. Mol. Biol. (2007) xx, xxx–xxx

Available online at www.sciencedirect.com

FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP

Elaine Small[1]†, Rachel Marrington[3]†, Alison Rodger[3], David J. Scott[6] Katherine Sloan[1], David Roper[4], Timothy R. Dafforn[5] and Stephen G. Addinall[1,2]*

[1]*Faculty of Life Sciences University of Manchester Michael Smith Building Oxford Road, Manchester M13 9PT, UK*

[2]*CISBAN, School of Clinical Medical Sciences–Gerontology University of Newcastle Henry Wellcome Laboratory for Biogerontology Research Newcastle General Hospital Newcastle upon Tyne NE4 6BE UK*

[3]*Department of Chemistry University of Warwick Coventry, CV4 7AL, UK*

[4]*Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK*

[5]*Department of Biosciences University of Birmingham Birmingham B15 2TT, UK*

[6]*National Centre for Macromolecular Hydrodynamics, School of Biosciences, University of Nottingham, Sutton Bonington Leicestershire LE12 5RD, UK*

**Corresponding author*

Cell division is a fundamental process for both eukaryotic and prokaryotic cells. In bacteria, cell division is driven by a dynamic, ring-shaped, cytoskeletal element (the Z-ring) made up of polymers of the tubulin-like protein FtsZ. It is thought that lateral associations between FtsZ polymers are important for function of the Z-ring *in vivo*, and that these interactions are regulated by accessory cell division proteins such as ZipA, EzrA and ZapA. We demonstrate that the putative *Escherichia coli* ZapA orthologue, YgfE, exists in a dimer/tetramer equilibrium in solution, binds to FtsZ polymers, strongly promotes FtsZ polymer bundling and is a potent inhibitor of the FtsZ GTPase activity. We use linear dichroism, a technique that allows structure analysis of molecules within linear polymers, to reveal a specific conformational change in GTP bound to FtsZ polymers, upon bundling by YgfE. We show that the consequences of FtsZ polymer bundling by YgfE and divalent cations are very similar in terms of GTPase activity, bundle morphology and GTP orientation and therefore propose that this conformational change in bound GTP reveals a general mechanism of FtsZ bundling.

Keywords: FtsZ; tubulin; polymer; bundling; GTP

Introduction

Control of cell division is of fundamental importance for both eukaryotic and prokaryotic organisms. In each case, cytoskeletal proteins play a role in this control. In bacteria, the main cytoskeletal element involved in cell division is the Z-ring, a membrane-associated structure consisting of polymers of the essential tubulin-related protein FtsZ.[1,2] The centrally

† E.S. and R.M. contributed equally to this work.

Abbreviations used: LD, linear dichroism; CD, circular dichroism.

E-mail address of the corresponding author: s.g.addinall@ncl.ac.uk

0022-2836/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.

located Z-ring is probably formed by bi-directional polymerization from a single nucleation site.[3,4] The Z-ring then reduces in diameter during cell division, drawing the cytoplasmic membrane inwards.[5,6] A number of accessory proteins that are essential for cell division require to be located to the cell centre during cell division.[7] These include membrane-anchored proteins protruding into the cytoplasm (ZipA, FtsK) or the periplasm (FtsQ, FtsI, FtsL, FtsN, YgbQ), multiple membrane spanning proteins (FtsW) and cytoplasmic proteins (FtsA). The localization of these proteins is achieved through recruitment to the Z-ring, by either direct or indirect interaction with FtsZ.[7] Another group of proteins, MinC, ZapA, EzrA and SulA, interact directly with FtsZ and influence its *in vivo* function,[8–13] but are not normally essential for cell survival.

FtsZ is a GTPase that polymerizes in a GTP-dependent fashion into linear, unbranched protofilaments.[2,14] Polymerization is co-operative under most conditions and, once formed, the polymers remain in a dynamic steady state until GTP availability becomes limiting.[15] As GTP is hydrolysed and the GDP:GTP ratio increases in solution,[16,17] GDP bound to FtsZ promotes bending of the polymers and de-polymerization.[18] Lateral associations between FtsZ protofilaments are observed *in vitro* to a greater or lesser degree depending on experimental conditions. Thick filaments, consisting of two parallel protofilaments, represent the simplest form of lateral association.[19–21] Higher-order lateral associations such as sheets (or ribbons), tubes and bundles[13,19,21–26] are formed from thick filaments in either parallel or antiparallel arrangements.[19,21,22]

With the exception of FtsA, there is evidence that each of the accessory cell division proteins that interact directly with FtsZ *in vivo* (ZipA, SulA, MinC, EzrA and ZapA), influence the biochemical properties of FtsZ *in vitro*.[9,11–13,25,27] These proteins can be conveniently grouped into inhibitors (SulA, MinC, EzrA) or supporters (ZipA, ZapA) of FtsZ polymerization. SulA expression is induced in response to DNA damage and it acts to depolymerize Z-rings, thus preventing cell division until DNA can be repaired.[8,28] MinC and its binding partner MinD are expressed constitutively. Together they act to prevent formation of Z-rings at cell poles, thus helping to ensure that cells divide precisely at their centre.[29] EzrA is thought to prevent aberrant Z-ring formation by inhibiting FtsZ assembly.[10] *In vitro* studies have demonstrated that SulA and MinC act to depolymerize FtsZ polymers,[9,12] whereas EzrA prevents FtsZ polymerization, possibly by inhibiting lateral associations between FtsZ protofilaments.[11] Strangely, SulA is the only one of these proteins reported to inhibit the FtsZ GTPase activity. Less is known about the *in vitro* properties of promoters of FtsZ polymerization. As yet, there are no candidates for the theoretical nucleation signal, which provides the starting point for polymerization of the Z-ring. However, both ZipA and ZapA are reported to stabilize Z-rings *in vivo*[13,30,31] and to bundle FtsZ polymers *in vitro*.[13,25,27]

We are interested in promoters of FtsZ polymerization, since it seems increasingly likely that the formation of bundles through lateral interactions between FtsZ polymers is critical to the function of the Z-ring. Using a combination of linear dichroism (LD) and circular dichroism (CD) we identified a specific molecular rearrangement associated with FtsZ polymer bundling induced by divalent cations.[32] LD is the measure of the difference in absorbance by a sample of light polarized in orthogonal directions; and in order to obtain an LD signal, molecules must be aligned with respect to a known orientation. We have achieved this by the use of a Couette flow cell,[33] which induces alignment as a result of shear flow in a liquid. Molecules that do not align (monomers or short oligomers), absorb the two polarizations of incident light equally and so do not give a signal in LD analysis.[34] This means that the spectra obtained in LD experiments are specifically from molecules in, or bound to, linear polymers.[34–36] We demonstrated that bundling caused by calcium (and some other divalent cations at appropriate concentrations) leads to significant, reproducible and characteristic changes in the LD absorbance spectrum of FtsZ polymers, that are not detectable by CD.[32] Specifically, an increase in LD signal intensity observed at 210–222 nm, due to increased polymer rigidity and/or length and an inversion of signal (from negative to positive) in a band at 250 nm, diagnostic of a dramatic shift in the orientation of the guanine moiety within the polymer. During this re-orientation, the long axis of the guanine group rotated at least 20° with respect to the axis of the FtsZ polymers, whereas the short axis showed no detectable movement. We therefore coined the term "flipping" to describe the guanine re-orientation. We proposed that such a significant movement of the guanine moiety would result in GTP adopting a sub-optimal conformation for hydrolysis and therefore could account for the reduction in FtsZ GTPase activity and increased stability observed in FtsZ polymer bundling.[32] However, a cause and effect relationship between guanine flipping and reduced GTPase activity has yet to be demonstrated.

ZapA orthologues from *Bacillus subtilis* and *Pseudomonas aeruginosa* have been studied previously.[13,37] The *B. subtilis* orthologue is thought to play a role in regulating Z-ring formation in conjunction with EzrA, and was shown to promote FtsZ polymer bundling.[13] The crystal structure of the *P. aeruginosa* ZapA orthologue has been determined, showing the protein as a dimer or tetramer: monomers and dimers associating through their C-terminal coiled-coil regions.[37] How these proteins interact with FtsZ and promote polymer association remains uncharacterised however. We chose to investigate the putative ZapA orthologue in *E. coli* which, compared to the *B. subtilis* and *P. aeruginosa* ZapA orthologues, is 11% and 25% identical in amino acid sequence, respectively (31% and 49% similar). It should be noted that *E. coli* does not possess an EzrA orthologue detectable by sequence analysis.

Using purified His-tagged YgfE and purified FtsZ, we have established that YgfE binds to and bundles

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

FtsZ polymers and is therefore a genuine functional homologue of ZapA. We also find that wild-type YgfE is a potent inhibitor of the FtsZ GTPase activity. We investigate the stoichiometry of the YgfE:FtsZ interaction and show that YgfE occurs as a dimer/tetramer equilibrium in solution. Using LD, we demonstrate that YgfE-mediated FtsZ polymer bundling induces a conformational change in the guanine moiety that is very similar to that induced by calcium and other bundling agents. On the basis of these results we are able to present testable models for YgfE-mediated polymer bundling that propose a role for guanine-flipping in this process.

Results

Polymeric structures formed by FtsZ *in vitro* have been described in a number of ways. As per Löwe and Amos,[19] we refer to a one molecule diameter, linear association of FtsZ monomers as a "protofilament" and a back-to-back, parallel association of two protofilaments as a "thick filament". All of our experiments are carried out under conditions that have previously been determined to support co-operative FtsZ polymerisation.[15] Since single protofilaments are not formed co-operatively,[20] we assume that the simplest newly formed polymers in our experiments are thick filaments. The words "sheet", "ribbon", "tube" and "tubule" refer to well defined flat or tubular arrangements of FtsZ protofilaments or thick filaments, which we do not observe under our polymerization conditions. We do observe "bundles," which we define as lateral associations between thick filaments and/or protofilaments that vary in their diameter and in their abundance depending on the polymerization conditions. It is not known whether protofilaments and thick filaments within bundles are parallel, antiparallel or a mixture of both.[37] We use the words "polymer" and "polymerization" as general terms to indicate multimeric FtsZ.

Purification and initial characterization of His_6YgfE

We sought to test whether YgfE, the putative *E. coli* orthologue of ZapA, has similar biochemical properties to ZapA. We engineered the expression of a six histidine-tagged YgfE protein (see Materials and Methods) which was purified to >90% purity using affinity chromatography (Supplementary Data Figure S1a). The circular dichroism spectrum of YgfE (Supplementary Data Fig. S1b) shows two minima at 208 and 222 nm, indicating the presence of significant alpha helical character. This is consistant with the published X-ray structure of ZapA from *P. aeruginosa*, which is dominated by a central alpha-helical coiled-coil domain.[13,37]

YgfE binds to FtsZ polymers and enhances their sedimentation

Using a sedimentation assay for polymerization of FtsZ we found that FtsZ pelleting in the presence of GTP (Figure 1, 2P) is enhanced by the presence of YgfE equimolar to FtsZ (Figure 1, 3P). Under these conditions, YgfE is found in the pellet with FtsZ polymers (Figure 1, 3P), whereas in the absence of FtsZ polymers there is no significant pelleting of YgfE (Figure 1, 4S and P). Consistent with published data,[15,24] the presence of 10 mM calcium enhances sedimentation of FtsZ polymers (Figure 1, 5P) and, under these conditions, YgfE again co-pellets with the FtsZ polymers (Figure 1, 6P). Hence YgfE co-sediments with FtsZ polymers in either the presence or absence of calcium and YgfE enhances the pelleting of FtsZ polymers in the absence of calcium.

YgfE enhances polymerization of FtsZ

We tested whether the binding of YgfE to FtsZ that we had detected (Figure 1) enhances FtsZ polymerization and/or stabilizes the polymers, as suggested by the increased pelleting of FtsZ in the presence of YgfE (Figure 1, compare lanes 3P and 2P). Using right-angled light scattering, we found that YgfE greatly enhances both the amount of polymer formation and the stability of FtsZ polymers (Figure 2). That is, the Δ of polymerization (Δ is defined in Materials and Methods) after addition of YgfE at 0.5 times the FtsZ concentration (Figure 2(a), open circles, YgfE added at arrow) was approximately 23-fold higher than the control value (Figure 2(a), filled circles and open circles before arrow). The $t^{0.5}$ of polymerization ($t^{0.5}$ is defined in Materials and Methods) was also increased by more than 8.5-fold (Figure 2(a), open circles). YgfE had similar affects

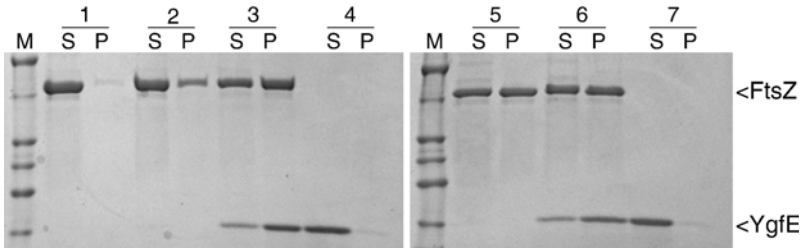


Figure 1. Purified, His-tagged YgfE binds to FtsZ. Supernatant (S) and pellet (P) lanes are marked for pelleting experiments (1 to 7). Where present, His_6-YgfE and FtsZ were at 8.3 μM in polymerization buffer (see Materials and Methods; YgfE concentrations are as that of the monomeric form throughout), GDP and GTP were at 2.5 mM and $CaCl_2$ was at 10 mM. (1) FtsZ + GDP, (2) FtsZ + GTP, (3) FtsZ + YgfE + GTP, (4) YgfE + GTP, (5) FtsZ + GTP + Ca^{2+}, (6) FtsZ + YgfE + GTP + Ca^{2+}, (7) YgfE + GTP + Ca^{2+}. Positions of FtsZ and YgfE proteins are marked.

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

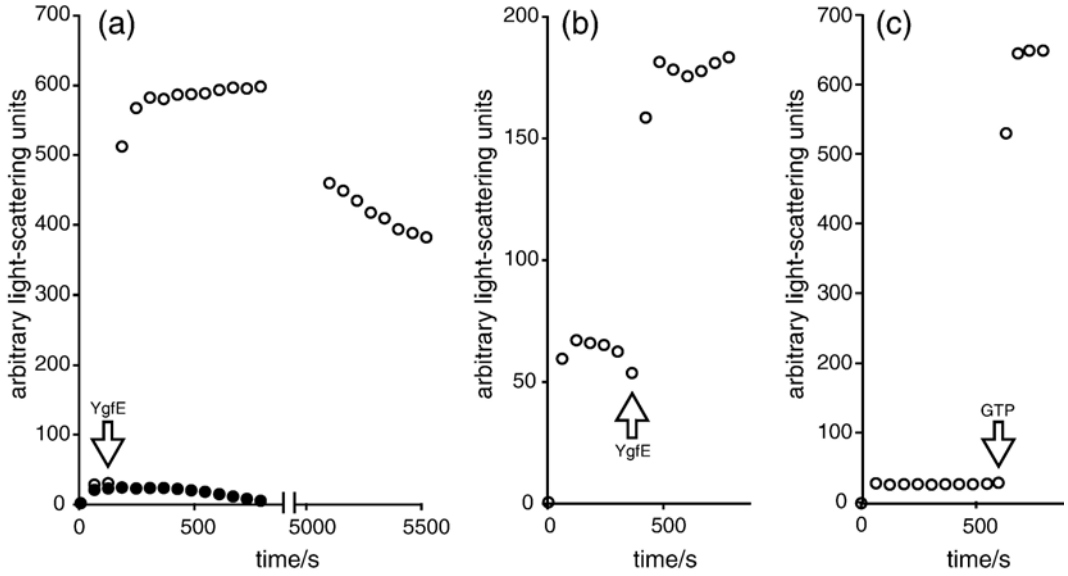


Figure 2. YgfE enhances polymerization of FtsZ. Polymerization of FtsZ was measured by right-angled light scattering. In each experiment FtsZ was at 8.3 μM in polymerization buffer and GTP added to 0.2 mM. For experiments (a), (b) and (c) the spectrofluorimeter voltage was set to 460, 500 and 470 mV, respectively. (a) GTP added at time = 0 s then 4.1 μM YgfE added at arrow (open circles). Control reaction had no YgfE added (filled circles). (b) GTP added at time = 0 s then 0.41 μM YgfE added at arrow. (c) YgfE added to 8.3 μM at time = 0 s then GTP added at arrow.

on FtsZ polymerization whether it was present before (Figure 2(c)) or added after GTP was introduced to start polymer formation (Figure 2(a) and (b)). Addition of YgfE at as little as 0.05 times the concentration of FtsZ also significantly enhanced both the Δ and the $t^{0.5}$ (Figure 2(b), YgfE added at arrow). Addition of YgfE to GDP-bound FtsZ caused a small increase in light scattering (Figure 2(c); YgfE added at zero time-point), which at least partially reflects the increased protein concentration in the sample (data not shown). Addition of GTP to this mixture, however, resulted in a 22-fold increase in light scattering (Figure 2(c); arrow).

YgfE induces bundling of FtsZ polymers

Increased longevity of FtsZ polymers, as we observed in the presence of YgfE (Figure 2), is usually associated with a reduction in GTPase activity and sometimes with polymer bundling.[13,15,38–40] We examined the morphology of FtsZ polymers formed in the absence and presence of YgfE by negative stain electron microscopy. In the absence of YgfE, FtsZ polymerized in a GTP-dependent fashion, into linear unbranched polymers (Figure 3(a)) with some tendency to form lateral associations (e.g. Figure 3(a), arrow) as previously reported.[23,40,41] In the presence of equimolar YgfE, similar polymers were formed, however, they readily associated in a lateral fashion into large bundles (Figure 3(b)). We were able to observe polymer bundling to a lesser degree when FtsZ was polymerized in the presence of YgfE at 0.05 times the FtsZ concentration; however, we were unable to detect any polymeric structures in experiments containing YgfE plus GTP, FtsZ plus GDP or YgfE plus GDP-bound FtsZ (data not shown).

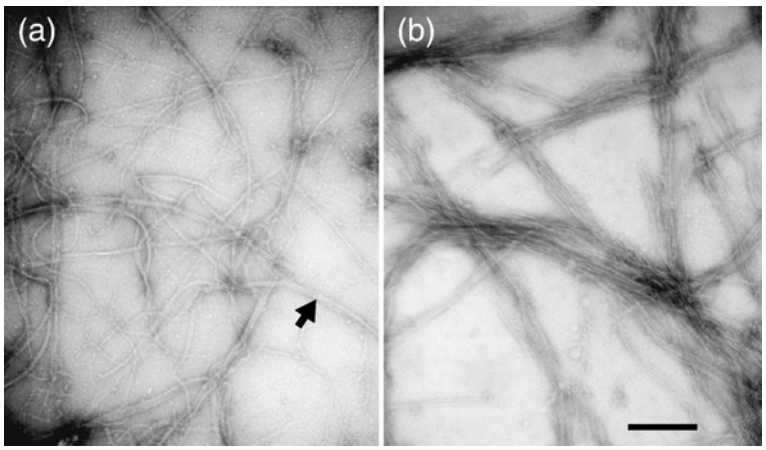


Figure 3. YgfE induces bundle-formation in FtsZ polymers. Negative stain electron microscopy of FtsZ polymers (8.3 μM) in the absence (a) or presence (b) of YgfE (8.3 μM). Bar represents 235 nm and applies to both (a) and (b). Arrow indicates example of lateral interactions between polymers in the absence of YgfE.

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

YgfE inhibits the FtsZ GTPase activity: titration of enhanced FtsZ polymer formation induced by YgfE

Having demonstrated that YgfE acts to promote sedimentation and bundling of FtsZ polymers, we determined the effect of YgfE on FtsZ GTPase activity. ZapA from *B. subtilis* has been reported[13] to have a minor effect on the FtsZ GTPase activity *in vitro* (~24% reduction when ZapA is equimolar to FtsZ). Our preliminary experiments indicated a strong inhibition of the GTPase activity of 8.3 μM FtsZ by 4.1 μM YgfE (Figure 2(a) and data not shown). We therefore titrated the effects of YgfE on these three aspects of FtsZ polymerization:GTPase activity, sedimentation and bundling.

Multiple measurements were made of each parameter with respect to YgfE concentration (Figure 4). In each case the concentration range for YgfE over which the effect takes place is similar, indicating that the phenomena are likely to be linked.

Analysis of YgfE-induced polymer bundling by linear dichroism

Having established that YgfE induces bundling of FtsZ polymers (Figure 3), we used LD to examine the structure of the bundles. The LD spectra of FtsZ polymers formed in the presence of YgfE showed clear differences from the LD spectra of FtsZ polymerized in the absence of YgfE (Figure 5). These differences were (i) a large increase in signal in the backbone region of the spectrum (210–222 nm, Figure 5(a)) indicative of increased polymer rigidity and/or length and (ii) a change in signal from negative to positive of a band at 250 nm (Figure 5(a) and (b)) indicating a large change in conformation of the guanine moiety (see Materials and Methods). These features were very similar to changes observed upon FtsZ polymer bundling induced by calcium[32] (Figure 5(a) and (b)) and other divalent cations.[32] We therefore conclude that the GTP in FtsZ polymers bundled by YgfE adopts the same characteristic conformation as in those bundled by calcium, indicating that a similar mechanism is involved in each case. Ruthenium red, a polycation which binds to tubulin, has been reported to bundle FtsZ polymers at sub-stoichiometric levels.[38] In that study, ruthenium red was shown to compete with calcium for binding to FtsZ and therefore is proposed to bind to the same site on FtsZ as calcium.[38] We found that ruthenium red induced qualitatively similar changes in the LD spectrum of FtsZ polymers to those induced by calcium and YgfE, but was a significantly less potent bundling agent than YgfE (Figure 5(a) and (b)).

We have previously demonstrated that the change in LD signal at 222 nm can be used to monitor FtsZ polymerization in a similar way to right-angled light scattering.[32] We therefore used this approach to determine the effect of different concentrations of YgfE on FtsZ polymer bundling. In each case where YgfE was present, we observed an increase in LD signal relative to the FtsZ only control (Figure 5(c)). In the presence of YgfE the increase in LD signal after addition of GTP appeared biphasic, with most of the very fast initial phase being during the 40 s dead time of the experiment (see Materials and Methods). A similar biphasic pattern was seen in light-scattering experiments in the presence of YgfE and the rate constant for the slower phase was found to be very similar when monitored by either technique (Figure 2(a) and data not shown). The biphasic nature of the increase in signal was most pronounced at higher YgfE concentrations and became less so as the YgfE concentration was reduced. The lowest YgfE concentration that maintained optimal bundle formation (maximum LD signal) resulted in a ratio of 1 YgfE to 2 FtsZ, consistent with titrations

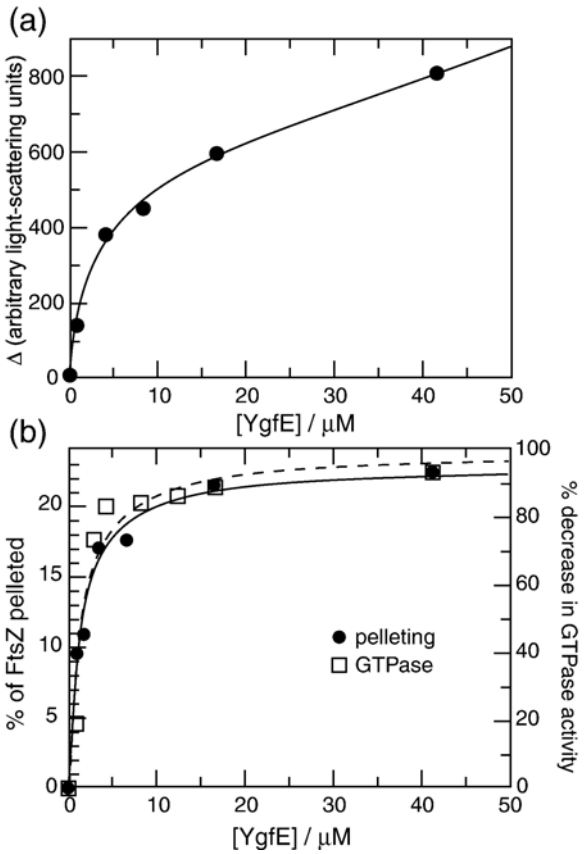


Figure 4. Titration of the effect of YgfE on FtsZ polymer sedimentation, light scattering and GTPase activity. FtsZ (8.3 μM) was polymerized by addition of GTP in the presence of varying concentrations of His-tagged YgfE. (a) The light scattering Δ of polymerization (with spectrofluorimeter voltage set to 460 mV) plotted against YgfE concentration. (b) FtsZ pelleting (expressed as percent decrease in amount of FtsZ remaining in supernatant due to addition of YgfE, relative to no YgfE control; filled circles) and reduction of FtsZ GTPase activity (open squares; values are the average of at least two independent experiments) plotted against YgfE concentration.

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

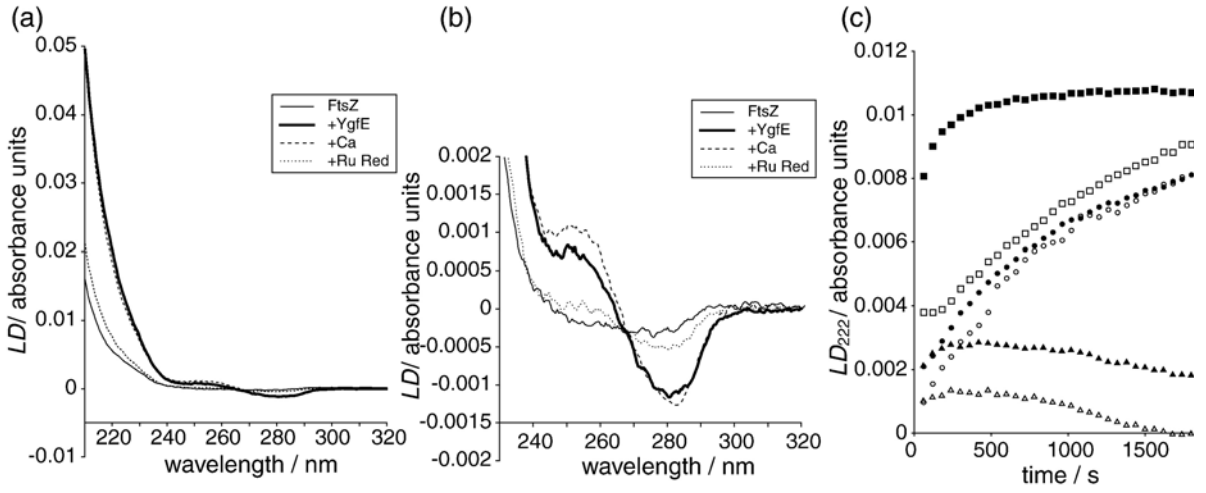


Figure 5. Analysis of YgfE-induced FtsZ polymer bundling by linear dichroism. (a) and (b) LD spectrum of FtsZ (7 μM) in polymerization buffer after addition of GTP (0.2 mM) in the absence (thin continuous line) and presence of different bundling agents: 3.5 μM YgfE (thick continuous line), 10 mM calcium (broken line) and 7 μM ruthenium red (dotted line). (a) Full spectrum, and (b) an expanded view of the near-UV region. (c) Time-course of LD signal at 222 nm, to monitor FtsZ polymer formation in the presence of increasing concentrations of YgfE. YgfE to FtsZ ratios of 0.1:1 (filled triangles), 0.5:1 (filled squares), 1:1 (open squares), 2:1 (filled circles) and 5:1 (open circles) were measured alongside an equivalent FtsZ only (8.3 μM) control (open triangles). GTP (0.25 mM) was added at time = 0 s.

of GTPase activity and FtsZ polymer formation (Figure 4).

Oligomerization of YgfE protein

We considered the possibility that YgfE might act as a multimer as suggested previously for ZapA from *B. subtilis*[13] and demonstrated in crystals of ZapA from *P. aeruginosa*.[37] Examination of purified His-tagged YgfE by dynamic light scattering predicted the molecular mass of individual particles in solution to be 58 kDa (data not shown). This corresponds closely to four times the predicted molecular mass of the His_6-YgfE protein (14.4 kDa), suggesting that YgfE may occur as a tetramer. To more accurately determine the stoichiometry of YgfE in solution, we performed a sedimentation equilibrium experiment by analytical ultracentrifugation. Data were collected from three loading concentrations (1000, 500 and 20 nM) brought to equilibrium at 15,000, 18,000 and 23,000 rpm (Figure 6). Analysis of the equilibrium data using a single species model gave a molecular mass of 53.9 (51.1–56.7) kDa, where the error is the 95% confidence limits from the global non-linear least squares regression. This is not that expected for a tetramer (57.6 kDa). Two mechanisms could yield an apparently reduced value of molecular mass: non-ideality and dissociation into smaller subunits. It is well known that non-ideality occurs above ca 3–4 mg/ml.[42] The second virial coefficient term (B22) only becomes significant at concentrations exceeding 3–4 mg/ml, and in fact can only be measured accurately at concentrations above 10 mg/ml. Given that the highest concentration used here was 0.1 mg/ml, non-ideality is not an issue. This leaves the second option: dissociation. Global analysis of the equilibrium traces in terms of a dimer/tetramer equilibrium yielded a dissociation constant of 320 [200,510] nM, where the figures in brackets are the 95% confidence limits from the non-linear global least squares analysis of nine data sets (Figure 6(a)). Given that the loading concentrations were in the order of 200–1000 nM, this dissociation constant is very well defined by this analysis, as appreciable amounts of both species are present (Figure 6(b)). Further analysis of the data in terms of higher-order associations (tetramer–octamer association) did not yield a significantly better fit, as judged from the statistics of the fit (data not shown), therefore we conclude that YgfE exists in dimer/tetramer equilibrium with a well defined submicromolar dissociation constant.

Discussion

A number of proteins are known to modulate cell division by interacting directly with FtsZ and changing its properties (ZipA, SulA, MinC, EzrA and ZapA). None of these proteins are completely essential for cell growth and division,[10,13,28,43] however, they all have interesting and different effects on FtsZ, which have been shown to be important for regulation of cell division. Previous data have shown that a green fluorescent protein (GFP)-fusion with the YgfE protein is localized to the Z-ring[13] and that, like *zapA* in *B. subtilis*, the *ygfE* gene is not essential for cell growth and division.[44] Never-the-less, YgfE-induced polymer bundling serves as a tractable and physiologically relevant system to study lateral interactions between FtsZ polymers (although it should be noted that our data refer to the His-tagged version of YgfE).

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

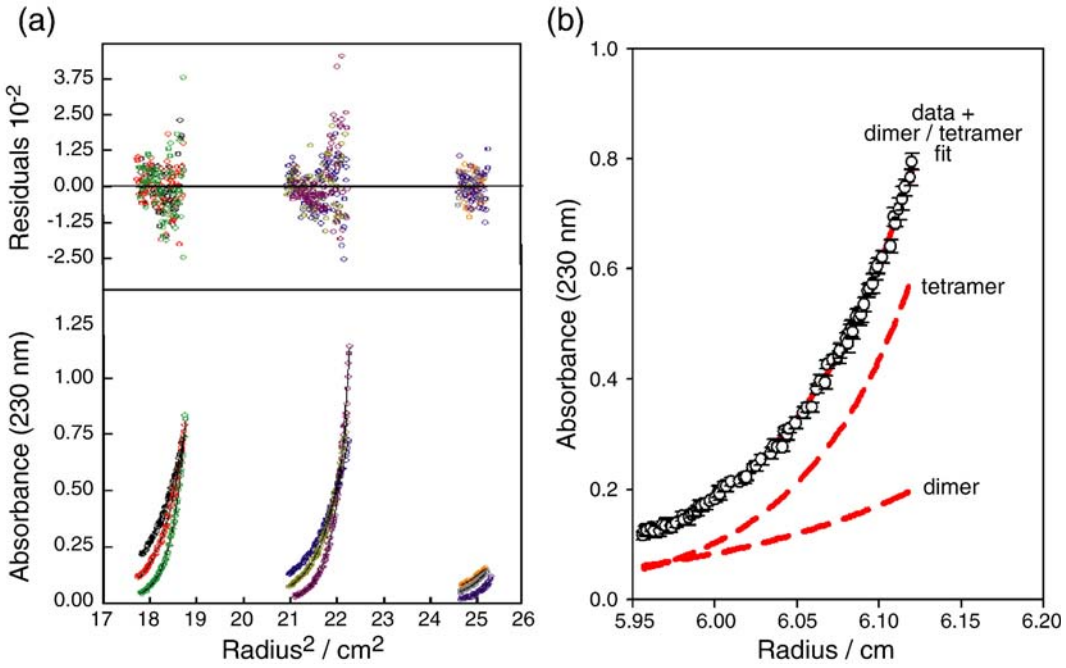


Figure 6. Sedimentation equilibrium of YgfE. (a) Data from three loading concentrations (1000, 500 and 200 nM w.r.t. monomers) is shown by symbols. The global fit to a dimer/tetramer model is shown by the continuous lines. As can be seen, the global fit to the model describes all nine data sets very well, as judged by the residuals of the fit displayed above the fitted data. The determined dissociation constant was 320 [200,510] nM (w.r.t. dimers), where the numbers in brackets are the 95% confidence limits from the global non-linear least squares analysis. (b) Deconvolution of data at 1 μM loading concentration brought to equilibrium at 18,000 rpm. The data are represented by the open circles and the error bars represent the 67% confidence limit in ten measurements. The continuous line through the data represents the global dimer-tetramer model fit. The broken lines represent the amount of dimeric and tetrameric YgfE present in the sample, as calculated from the global fit. As can be seen, an appreciable amount of both species are present, allowing for accurate determination of the dissociation constant.

We have demonstrated that the YgfE protein binds to FtsZ polymers in pelleting experiments and that, in the absence of calcium, YgfE promotes FtsZ sedimentation (Figure 1). When both calcium and YgfE are present together (Figure 1, compare lanes 6P and 5P) we do not observe an additive effect on FtsZ pelleting, suggesting that the two reagents mimic each others' actions and perhaps compete for binding to the FtsZ polymers. In LD experiments, we note that the combination of YgfE and Ca^{2+} exhibits intermediate effects upon FtsZ polymerization, compared to either in isolation (data not shown), again suggesting competition for interaction with FtsZ.

We have also shown that YgfE promotes polymerization of FtsZ, as monitored by right-angled light scattering (Figure 2) resulting in extensive polymer bundling (Figure 3). The interaction is saturable at sub-stoichiometric levels (Figure 4). Effects on FtsZ polymerization similar to these have been shown previously for ruthenium red[38] and ZapA.[13] In contrast, calcium[15,24] and glutamate[45] (and our unpublished results) induce similar increases in polymerization but at unphysiological 1000-fold and 6×10^7-fold excess concentrations, respectively.

Our results with YgfE are similar to those obtained previously with ZapA but with notable differences. Gueiros-Filho and Losick reported a comparable increase in light scattering upon addition of ZapA to FtsZ whether in the presence or absence of GTP.[13] In contrast, we find that the increase in FtsZ light scattering induced by YgfE is 22-fold greater in the presence of GTP than it is in the presence of GDP (Figure 2(c)). This indicates that YgfE does not induce monomeric FtsZ to form higher-order structures. Consistent with this, we did not detect oligomeric structures by electron microscopy when YgfE was incubated with FtsZ in the presence of GDP or with no nucleotide phosphate added. Indeed, we found that the light scattering increase observed under these conditions (Figure 2(c), before arrow) was at least partially accounted for by the addition of extra protein (data not shown). We therefore find by two different techniques (light scattering and electron microscopy) that YgfE has no significant effect on unpolymerised FtsZ.

We have shown that YgfE is a potent inhibitor of the GTPase activity of *E. coli* FtsZ (approximately 87% reduction when equimolar (Figure 4(b))). In contrast, ZapA was shown to have a smaller effect on the GTPase activity of *B. subtilis* FtsZ (approximately 24% reduction when equimolar[13]). These differences between our results with YgfE and previous data on *B. subtilis* ZapA could reflect either technical differences between laboratories or subtle differences between bacterial species.

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

We have recently reported that FtsZ polymers formed in the absence of YgfE have a reproducible LD spectrum[32] that is consistent with the FtsZ protofilament structure predicted by other methods.[19] We also reported that bundling induced by calcium and other divalent cations (Mg^{2+}, Ca^{2+} and Cu^{2+}) produces characteristic changes in this spectrum. These changes, namely an increase in absorbance at 210–222 nm and a change in shape from a negative maximum to a positive maximum at 250 nm, indicate a stiffening and/or lengthening of the polymers and a dramatic shift in the orientation of the guanine moiety of bound GTP, respectively.[32] We show here that polymerization of FtsZ in the presence of YgfE leads to these same changes in the LD spectrum (Figure 5(a) and (b)). This means that sub-stoichiometric levels of YgfE protein and 1000-fold excess of Ca^{2+} lead to very similar conformational changes in GTP bound to FtsZ polymers. That these two molecules (and ruthenium red, Figure 5(b)) result in re-orientation of the guanine moiety is also consistent with the strong inhibition of FtsZ GTPase activity exhibited by both (Mukherjee and Lutkenhaus;[15] and this study).

Bundling of cytoskeletal polymers by accessory proteins is a common biological phenomenon. In particular, there are many examples of actin-binding proteins that bind to and stabilize actin polymers, many of which have been demonstrated to induce bundling of actin fibres *in vitro*.[46,47] Even so, what we present here is, to our knowledge, the first demonstration of a conformational change of bound nucleotide within a protein polymer, resulting from accessory protein-induced bundling.

We proposed previously[32] that guanine-flipping induced by divalent cations moves the GTP from a conformation that is optimal for GTP hydrolysis to one that is suboptimal for hydrolysis. We also proposed that this movement could alter the surface properties of FtsZ polymers, such that lateral interactions between the polymers are favoured, resulting in bundling. YgfE induces a remarkably similar change in the orientation of the guanine moiety relative to the FtsZ polymer axis (Figure 5). Taking account of this, together with the adoption of a dimer/tetramer equilibrium by YgfE in solution (Figure 6), we can envisage two possible models for YgfE-mediated FtsZ polymer bundling where flipping represents an active part of the bundling process.

Model (1): YgfE binding to the FtsZ polymer induces guanine flipping, which in turn alters surface properties of the protofilament adjacent to the cleft occupied by GTP.[32] This configuration of the FtsZ polymer is inherently more likely to undergo lateral interactions and so bundling occurs. In order to induce guanine-flipping without demonstrably changing the conformation of FtsZ itself, we propose that YgfE is most likely to bind at or near the GTP binding pocket, changing the environment of the pocket such that the alternative GTP conformation is favoured. Model (2): The observation that YgfE is in a tetramer–dimer equilibrium in solution suggests that it has either a divalent or tetravalent binding mode, as also proposed for ZapA from *P. aeruginosa*.[37] That is, YgfE could simultaneously contact two or four FtsZ protofilaments or thick filaments. It follows from this that YgfE could act as a cross-bridge bringing two or more FtsZ polymers together. Subsequent to this initial cross-linking event, kinetically unfavourable lateral interactions could be stabilized by flipping of the guanine moiety leading to altered surface properties of the protofilament. More in-depth kinetic studies, to determine exactly when flipping occurs during the bundling process, will be required to distinguish between these two models.

A third hypothesis, that we do not favour, is that guanine-flipping is merely a secondary consequence of polymer bundling. That is, after the important molecular interactions between FtsZ and YgfE have led to bundle formation and a reduction in GTPase activity, the guanine group is left in an alternative orientation. Our current data do not rule this out, however, we find it unlikely that a large shift in the position of the guanine moiety has no role in the function of a GTPase such as FtsZ.

The interesting observation that the effect of YgfE on the FtsZ GTPase activity is a substoichiometric phenomenon (Figures 2 and 4) may indicate some co-operative process involving conformational changes in FtsZ molecules distant from the bound YgfE. This may originate from the structure of the FtsZ protofilament where each monomer forms part of two active sites, one with each of the $n+1$ and $n-1$ FtsZ subunits. Alternatively this stoichiometry might be explained by model (2), where one YgfE tetramer could make simultaneous contact with multiple FtsZ polymers/thick filaments. The latter possibility raises topological issues regarding the interaction between ZapA and FtsZ that have been discussed previously.[37]

It is possible that YgfE inhibits the GTP hydrolysis step directly (as we have proposed in model (1)), or indirectly (as in model (2)), since this has previously been determined to be the rate-limiting step for the FtsZ GTPase.[14] Alternatively YgfE could inhibit another process such as phosphate release, GTP binding or GDP release to such an extent that this step becomes rate-limiting.

A number of these aspects of YgfE action are difficult to consider in our models until other fundamental questions are answered, for example: where is the FtsZ binding site on YgfE?; where is the YgfE binding site on FtsZ?; how many molecules of FtsZ does a YgfE dimer bind simultaneously?; if YgfE forms cross-links, do these occur between protofilaments or between thick filaments? Work is now underway to determine the answers to these questions.

Previous published data reveal that GFP-YgfE localizes to the Z-ring[13] and that the *ygfE* gene is not essential for survival.[44] Our data have added significantly to the characterization of YgfE and we confidently conclude that YgfE is a functional homologue of ZapA. We therefore propose to refer

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

to YgfE as the ZapA orthologue of *E. coli* in future work.

Materials and Methods

DNA manipulations

Chromosomal DNA from strain MG1655 was prepared using standard techniques[48] and then amplified further using the Genomiphi kit as per manufacturers instructions (Amersham Biosciences UK, Little Chalfont). The *ygfE* gene was amplified from this template DNA using Pfu Turbo polymerase (Stratagene, La Jolla, USA) and forward and reverse primers (MWG Biotech, Ebersberg, Germany: ATCGAAGAATTCCATGTCTGCACAACCCGTCGATCTCC; ATCGTTAAGCTTCCACAGTAAACCGAAAAGTGTCATTCAAAG) engineered to contain sites for EcoRI and HindIII, respectively (underlined). The amplified product was cloned into the pETDuet-1 expression vector (Novagen, EMD Biosciences Inc., Darmstadt, Germany) using EcoRI and HindIII (New England Biolabs, Hitchin, UK) to engineer a fusion between the six-histidine tag and the *ygfE* open reading frame. The resultant plasmid, pETDuet-ygfE, was transformed into competent BL21-AI cells (Invitrogen, Paisley, UK) and, upon induction with 0.2% (w/v) arabinose, was shown to over-produce a protein of the predicted size (14.4 kDa; data not shown).

Protein purification

FtsZ was purified as previously described, using ammonium sulfate precipitation, followed by ion -xchange chromatography.[16] Light scattering, pelleting, CD, and LD titration data were obtained using an FtsZ preparation with a GTPase activity of 1.8 mol/mol per min. In order to ascertain that LD data were consistent between FtsZ preparations, we obtained LD spectra from a second preparation with a GTPase activity of 3.5 mol/mol per min. Similar spectra were obtained with both preparations.

His-tagged YgfE was produced and purified by affinity chromatography as follows. A colony of BL21-AI pETDuet-ygfE was inoculated into 200 ml of LB containing ampicillin (50 $\mu g ml^{-1}$) and 0.1% (w/v) glucose, then shaken for 6.5 h at 37 °C. Arabinose was added to a final concentration of 0.2% and then the culture was left shaking at 37 °C for 16 h. Cells were harvested by centrifugation at 3542*g* in a 4K15 Benchtop centrifuge (Sigma, Osterode am Harz, Germany) for 20 min at 15 °C. The pellet was resuspended in 10 ml of CelLytic B solution (Sigma Aldrich, Poole, UK; containing 500 mM NaCl and protease inhibitors (Sigma Aldrich)). The lysate was cleared by centrifugation at 25,000*g* at 15 °C using a Beckman J2-HS centrifuge and JA25-50 rotor (Beckman Coulter (UK) Ltd., High Wycombe), filtered through a 0.22 μm disc filter (Millipore (UK) Ltd., Watford). The lysate was passed through a 1 ml His-Trap column, which was washed extensively and then the fusion protein was eluted with a step gradient of imidazole concentrations as per manufacturers instructions (Amersham Biosciences UK). The majority of His-YgfE protein was eluted at a concentration of 300 mM imidazole. Finally, this eluate was dialysed into storage buffer (50 mM Hepes-NaOH (pH 7.2), 0.1 mM EDTA, 10% (v/v) glycerol). We found that His-YgfE could be stored for long periods of time at room temperature (20 °C), 4 °C, −20 °C or −80 °C, with small amounts of precipitation occurring at the higher temperatures. His-YgfE concentration is given as that of the monomeric form throughout.

Pelleting assay

A protocol for pelleting FtsZ polymers similar to that outlined previously[49] was used. For initial experiments a 125 μl reaction containing FtsZ and/or YgfE was made up in polymerization buffer (50 mM Mes (pH 6.5), 10 mM $MgCl_2$, 50 mM KCl). Calcium chloride was added to 10 mM, where specified, and GTP or GDP were added to 2.5 mM, and the reaction mixed, immediately before centrifugation; a consistent dead time of less than 1 min was acheived. Reactions were centrifuged for 10 min, at 20 °C, in a TLA-100 rotor, in an Optima TLX centrifuge (Beckman Coulter (UK) Ltd) at 80,000 rpm. 100 μl of supernatant was collected, then the remainder was carefully removed from the pellet which was then resuspended in 125 μl of polymerization buffer. Equivalent volumes of supernatant and pellet were examined by SDS-PAGE, on a pre-cast 4–15% gradient gel (BioRad, Hemel Hempstead, UK). Titration experiments were performed in identical fashion except using a reaction volume of 50 μl and running on 8–16% gradient gels. The amount of FtsZ remaining in the supernatant when FtsZ was pelleted in the presence of GTP but the absence of YgfE was set arbitrarily to 100%. Increased pelleting due to addition of YgfE resulted in less FtsZ remaining in the supernatant (increased pelleting) and was expressed as a percentage of the control value.

GTPase assay

GTPase activity of newly purified FtsZ protein was determined using γ-^{32}P-labelled GTP (Amersham Biosciences, UK) and thin-layer chromatography as described.[16] This technique was also used for preliminary testing of the FtsZ GTPase activity in the presence of YgfE. For titration of the effects of YgfE on FtsZ GTPase activity, we measured production of NADH from an enzyme-linked assay which has been used previously with FtsZ.[50] This is an ATP production assay, which gave comparable values to the FtsZ GTPase activity determined by the radioactive method, but was more appropriate for multiple repeat experiments.

Electron microscopy

Negative stain electron microscopy was performed as described.[16]

Analysis of titration experiments

Data from measurements of sedimentation and GTPase activity, as a function of YgfE concentration, were fitted to a weak binding isotherm:

$$Y = \frac{L^{*}M}{K_d + L} + \mathrm{OFF} \quad (1)$$

where L is the YgfE concentration, M is the maximum signal change and K_d is the apparent binding constant and OFF is a baseline offset. Data from measurements of increase in light scattering resulting from bundling, as a

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

function of YgfE concentration, were fitted to a modified equation:

$$Y = \frac{L^*M}{K_d + L} + \mathrm{OFF} + L^*\mathrm{SLOPE} \quad (2)$$

where a term L*SLOPE is included to take into account the linear increase in light scattering, independent of bundling, caused by the introduction of YgfE.

Spectroscopic measurements

FtsZ polymerization assays were performed as described.[16] Briefly, FtsZ was incubated in polymerization buffer and GTP was added to the appropriate concentration to initiate polymerization. Where applicable, $CaCl_2$ was added to the polymerization buffer to 10 mM.

Right-angled light scattering was carried out as described,[16] except that in some experiments the voltage setting of the spectrofluorimeter was reduced from 500 V to 460 or 470 V in order to reduce sensitivity. This is because bundling of FtsZ by YgfE results in a huge increase in light scattering. As previously[16], light scattering traces are characterized by a combination of the maximal increase in light scattering upon addition of GTP, the delta (Δ), and the time to return to half maximal light scattering signal, $t^{0.5}$.

CD measurements were performed using a JASCO (Tokyo, Japan) J-810 spectropolarimeter at room temperature using a 0.5 mm pathlength demountable quartz cuvette. Data were collected every 0.2 nm from 260–190 nm, with a scanning speed of 10 nm min^{-1}, a response time of 0.5 s and averaging over 16 scans.

LD measurements were performed at room temperature using a JASCO J-715 spectropolarimeter adapted for LD spectroscopy.[33] Samples were aligned in the light beam using a custom made Couette cell (Crystal Precision Optics, Rugby, UK), which consists of a cylindrical cross-section sleeve with a removable quartz capillary (sealed at one end with Araldite Rapid®) held centrally with respect to its circular face by a rubber O-ring. A quartz rod is suspended from the demountable lid into the capillary, creating an annular gap between rod and inner capillary wall of ~0.25 mm. The rod and the centre of the capillary are aligned so that the capillary is able to rotate freely. The sample (25–40 μl) is placed into the capillary before loading into the cylindrical sleeve. The lid and rod unit are lowered into position and rotation of the capillary induces a shear force across the sample. The cylindrical sleeve has two windows for the light to pass through to make contact with the sample. The voltage applied to the motor that rotates the outer quartz cylinder is controlled electronically to allow the sample solution to be maintained with the highest possible degree of alignment without inducing turbulent flow and Taylor vortices. Data were collected using an interval scan measurement program that was available within the Jasco software. This enabled single full wavelength scans from 350–190 nm to be collected every minute at a scanning speed of 200 nm min^{-1}, data pitch 0.5 nm and with a response of 0.5 s and thus the monitoring of the kinetics of polymerisation across the whole wavelength spectrum. The spectra are of suitable intensity that only one scan was necessary. Baselines of data collected after de-polymerisation of FtsZ were subtracted from all spectra. The wavelength range of the LD spectra were in some cases restricted by the high absorbance of the samples at low wavelength. Data reported here are truncated at a point where the Beer-Lambert law is valid.

Interpretation of LD spectra

We used LD to look for changes in the orientations of FtsZ subunits or associated molecules during bundling. The only change in shape of the LD spectrum (indicating an orientation change) occurred in a band at approximately 250 nm, while a band of similar intensity at approximately 280 nm remained unchanged. (Changes at shorter wavelengths did not change the shape of the spectrum, just the intensity of the signal.) These two bands are very close to the two maxima of absorbance of guanine in this region of the spectrum (248 and 275 nm). There is no possible confusion from tryptophan absorbance (at 280 nm), since it is absent from both FtsZ and YgfE. The only other potential alternatives are phenylalanine (13 in FtsZ, absorbs at 257 nm) and tyrosine (three in FtsZ, absorbs at 274 nm), which we have ruled out for the following reasons: (a) in order for 13 Phe residues to have an additive absorbance signal they would have to be similarly oriented to the polymer axis (i.e. parallel with one another); the same can be said of the three Tyr residues: examination of published FtsZ structures suggests that these residues are essentially randomly oriented relative to the fibre axis and would consequently give very low LD signal if any; (b) 13 Phe residues, even if all in the same orientation, would absorb approximately as much as one guanine group; (c) three Tyr residues, even if all in the same orientation, would absorb approximately one third as much as one guanine group; (d) the ~250 nm band shows a significant change in orientation with respect to the fibre axis upon polymer bundling, which if representing movement of multiple Phe residues would be expected to have gross effects on the protein secondary structure and register in our previous CD experiments;[32] (e) the peak absorbance in both Ca and YgfE-bundled polymers appears to be closer to 250 nm than 255 nm; (f) when guanine is perpendicular to the fibre axis the ~250 nm and ~280 nm bands are of a similar magnitude in the LD signal. Since structural studies[19,51] indicate that GTP is regularly arranged in the protein polymer, thus occupying a consistent orientation with respect to the fibre axis, the overwhelming balance of probability indicates that the 250 nm and 280 nm bands originate from the strong guanine chromophore in GTP.

Dynamic light scattering

Dynamic light scattering measurements were carried out on a Dynapro. A 69 μM (1 mg/ml) sample of YgfE was filtered through a 0.3 μm filter to remove aggregated species. The sample was then introduced into a 0.2 cm × 0.2 cm pathlength quartz cuvette and the variation in light scattering with time measured over 30 min allowing >20 independent determinations of molecular size to be made. These data were deconvolved assuming a two component system (one component being YgfE and the second accounting for any aggregated material that is present in the sample) model as per the Dynapro analysis software.

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed on a Beckman Optima XL-A analytical ultracentrifuge using an AnTi 60 rotor. Data were obtained at 15,000, 18,000 and 23,000 rpm in cells with six channel centre pieces, with loading concentrations of 1000, 500 and 200 nM (w.r.t. monomers) YgfE. The absorbance

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

wavelength was set to 230 nm, and the temperature was 20 °C. Scans were collected after 12 h, then after two subsequent 4 h intervals. Scans were then deducted from each other and equilibrium judged to have been attained by the absence of systematic deviation in absorbance across the cell. The solution density of the solvent was measured using an Anton Paar DMA 5000 density meter, at a set temperature of 20 °C. The partial specific volumes of the samples were calculated and corrected to 20 °C using program SEDNTERP according to the amino acid sequence of the polypeptides.[52] The absorbance coefficient at 230 nm was calculated to be 360,000 M^{-1} cm^{-1}. Molecular masses and analysis of self-association were determined by fitting nine data sets (three loading concentrations at three speeds) using global non-linear least squares analysis using the program WINONLIN.[53] Errors were calculated at the 95% confidence level.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.03.025

References

1. Addinall, S. G. & Holland, B. (2002). The tubulin ancestor, FtsZ, draughtsman, designer and driving force for bacterial cytokinesis. *J. Mol. Biol.* **318**, 219–236.
2. Romberg, L. & Levin, P. A. (2003). Assembly dynamics of the bacterial cell division protein FtsZ: poised at the edge of stability. *Annu. Rev. Microbiol.* **57**, 125–154.
3. Addinall, S. G. & Lutkenhaus, J. (1996). FtsZ-spirals and -arcs determine the shape of the invaginating septa in some mutants of *Escherichia coli*. *Mol. Microbiol.* **22**, 231–237.
4. Ma, X., Ehrhardt, D. W. & Margolin, W. (1996). Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein. *Proc. Natl Acad. Sci. USA*, **93**, 12998–13003.
5. Bi, E. F. & Lutkenhaus, J. (1991). FtsZ ring structure associated with division in *Escherichia coli*. *Nature*, **354**, 161–164.
6. Addinall, S. G., Bi, E. & Lutkenhaus, J. (1996). FtsZ ring formation in fts mutants. *J. Bacteriol.* **178**, 3877–3884.
7. Vicente, M., Rico, A. I., Martinez-Arteaga, R. & Mingorance, J. (2006). Septum enlightenment: assembly of bacterial division proteins. *J. Bacteriol.* **188**, 19–27.
8. Bi, E. & Lutkenhaus, J. (1993). Cell division inhibitors SulA and MinCD prevent formation of the FtsZ ring. *J. Bacteriol.* **175**, 1118–1125.
9. Hu, Z., Mukherjee, A., Pichoff, S. & Lutkenhaus, J. (1999). The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. *Proc. Natl Acad. Sci. USA*, **96**, 14819–14824.
10. Levin, P. A., Kurtser, I. G. & Grossman, A. D. (1999). Identification and characterization of a negative regulator of FtsZ ring formation in *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA*, **96**, 9642–9647.
11. Haeusser, D. P., Schwartz, R. L., Smith, A. M., Oates, M. E. & Levin, P. A. (2004). EzrA prevents aberrant cell division by modulating assembly of the cytoskeletal protein FtsZ. *Mol. Microbiol.* **52**, 801–814.
12. Mukherjee, A., Cao, C. & Lutkenhaus, J. (1998). Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **95**, 2885–2890.
13. Gueiros-Filho, F. J. & Losick, R. (2002). A widely conserved bacterial cell division protein that promotes assembly of the tubulin-like protein FtsZ. *Genes Dev.* **16**, 2544–2556.
14. Romberg, L. & Mitchison, T. J. (2004). Rate-limiting guanosine 5′-triphosphate hydrolysis during nucleotide turnover by FtsZ, a prokaryotic tubulin homologue involved in bacterial cell division. *Biochemistry*, **43**, 282–288.
15. Mukherjee, A. & Lutkenhaus, J. (1999). Analysis of FtsZ assembly by light scattering and determination of the role of divalent metal cations. *J. Bacteriol.* **181**, 823–832.
16. Small, E. & Addinall, S. G. (2003). Dynamic FtsZ polymerization is sensitive to the GTP to GDP ratio and can be maintained at steady state using a GTP-regeneration system. *Microbiology*, **149**, 2235–2242.
17. Scheffers, D. J., den Blaauwen, T. & Driessen, A. J. (2000). Non-hydrolysable GTP-gamma-S stabilizes the FtsZ polymer in a GDP-bound state. *Mol. Microbiol.* **35**, 1211–1219.
18. Lu, C., Reedy, M. & Erickson, H. P. (2000). Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. *J. Bacteriol.* **182**, 164–170.
19. Löwe, J. & Amos, L. A. (1999). Tubulin-like protofilaments in Ca2+-induced FtsZ sheets. *EMBO J.* **18**, 2364–2371.
20. Romberg, L., Simon, M. & Erickson, H. P. (2001). Polymerization of FtsZ, a bacterial homolog of tubulin - is assembly cooperative? *J. Biol. Chem.* **276**, 11743–11753.
21. Löwe, J. & Amos, L. A. (2000). Helical tubes of FtsZ from *Methanococcus jannaschii*. *Biol. Chem.* **381**, 993–999.
22. Oliva, M. A., Huecas, S., Palacios, J. M., Martin-Benito, J., Valpuesta, J. M. & Andreu, J. M. (2003). Assembly of archaeal cell division protein FtsZ and a GTPase-inactive mutant into double-stranded filaments. *J. Biol. Chem.* **278**, 33562–33570.
23. Erickson, H. P., Taylor, D. W., Taylor, K. A. & Bramhill, D. (1996). Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers. *Proc. Natl Acad. Sci. USA*, **93**, 519–523.
24. Yu, X. C. & Margolin, W. (1997). Ca2+-mediated GTP-dependent dynamic assembly of bacterial cell division protein FtsZ into asters and polymer networks *in vitro*. *EMBO J.* **16**, 5455–5463.
25. Hale, C. A., Rhee, A. C. & de Boer, P. A. (2000). ZipA-induced bundling of FtsZ polymers mediated by an interaction between C-terminal domains. *J. Bacteriol.* **182**, 5153–5166.
26. Gonzalez, J. M., Jimenez, M., Velez, M., Mingorance, J., Andreu, J. M., Vicente, M. & Rivas, G. (2003). Essential cell division protein FtsZ assembles into one monomer-thick ribbons under conditions resembling the crowded intracellular environment. *J. Biol. Chem.* **278**, 37664–37671.
27. RayChaudhuri, D. (1999). ZipA is a MAP-Tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division. *EMBO J.* **18**, 2372–2383.
28. D'Ari, R. & Huisman, O. (1983). Novel mechanism of

Please cite this article as: Small, E. *et al.*, FtsZ Polymer-bundling by the *Escherichia coli* ZapA Orthologue, YgfE, Involves a Conformational Change in Bound GTP, *J. Mol. Biol.* (2007), doi:10.1016/j.jmb.2007.03.025

cell division inhibition associated with the SOS response in *Escherichia coli*. *J. Bacteriol.* **156**, 243–250.
29. Margolin, W. (2001). Bacterial cell division: a moving MinE sweeper boggles the MinD. *Curr. Biol.* **11**, R395–R398.
30. Liu, Z., Mukherjee, A. & Lutkenhaus, J. (1999). Recruitment of ZipA to the division site by interaction with FtsZ. *Mol. Microbiol.* **31**, 1853–1861.
31. Hale, C. A. & de Boer, P. A. (1999). Recruitment of ZipA to the septal ring of *Escherichia coli* is dependent on FtsZ and independent of FtsA. *J. Bacteriol.* **181**, 167–176.
32. Marrington, R., Small, E., Rodger, A., Dafforn, T. R. & Addinall, S. G. (2004). FtsZ fiber bundling is triggered by a conformational change in bound GTP. *J. Biol. Chem.* **279**, 48821–48829.
33. Marrington, R., Dafforn, T. R., Halsall, D. J. & Rodger, A. (2004). Micro volume Couette flow sample orientation for absorbance and fluorescence linear dichroism. *Biophys. J.* **87**, 2002–2012.
34. Rodger, A. & Nordén, B. (1987). *Circular and Linear Dichroism.* Oxford University Press, Oxford.
35. Rodger, A., Rajendra, J., Marrington, R., Ardhammar, M., Nordén, B., Hirst, J. B. *et al.* (2002). Flow oriented linear dichroism to probe protein orientation in membrane environments. *Phys. Chem. Chem. Phys.* **4**, 4501-4057.
36. Dafforn, T. R., Rajendra, J., Halsall, D. J., Serpell, L. C. & Rodger, A. (2004). Protein fiber linear dichroism for structure determination and kinetics in a low-volume, low-wavelength Couette flow cell. *Biophys. J.* **86**, 404–410.
37. Low, H. H., Moncrieffe, M. C. & Löwe, J. (2004). The crystal structure of ZapA and its modulation of FtsZ polymerisation. *J. Mol. Biol.* **341**, 839–852.
38. Santra, M. K., Beuria, T. K., Banerjee, A. & Panda, D. (2004). Ruthenium red induced bundling of bacterial cell division protein, FtsZ. *J. Biol. Chem.* **279**, 25959–25965.
39. White, E. L., Ross, L. J., Reynolds, R. C., Seitz, L. E., Moore, G. D. & Borhani, D. W. (2000). Slow polymerization of Mycobacterium tuberculosis FtsZ. *J. Bacteriol.* **182**, 4028–4034.
40. Mukherjee, A. & Lutkenhaus, J. (1998). Dynamic assembly of FtsZ regulated by GTP hydrolysis. *EMBO J.* **17**, 462–469.
41. Mukherjee, A. & Lutkenhaus, J. (1994). Guanine nucleotide-dependent assembly of FtsZ into filaments. *J. Bacteriol.* **176**, 2754–2758.
42. Wills, P. R. & Winzor, D. J. (2005). Allowances for thermodynamic non-ideality. In *Analytical Ultracentrifugation: Techniques and Methods* (Scott, D. J., Harding, S. E. & Rowe, A. J., eds), Royal Society of Chemistry, Cambridge, UK.
43. Geissler, B., Elraheb, D. & Margolin, W. (2003). A gain-of-function mutation in ftsA bypasses the requirement for the essential cell division gene zipA in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **100**, 4197–4202.
44. Johnson, J. E., Lackner, L. L., Hale, C. A. & de Boer, P. A. (2004). ZipA is required for targeting of DMinC/DicB, but not DMinC/MinD, complexes to septal ring assemblies in *Escherichia coli*. *J. Bacteriol.* **186**, 2418–2429.
45. Beuria, T. K., Krishnakumar, S. S., Sahar, S., Singh, N., Gupta, K., Meshram, M. & Panda, D. (2003). Glutamate-induced assembly of bacterial cell division protein FtsZ. *J. Biol. Chem.* **278**, 3735–3741.
46. Roger, B., Al-Bassam, J., Dehmelt, L., Milligan, R. A. & Halpain, S. (2004). MAP2c, but not tau, binds and bundles F-actin via its microtubule binding domain. *Curr. Biol.* **14**, 363–371.
47. Volkmann, N., DeRosier, D., Matsudaira, P. & Hanein, D. (2001). An atomic model of actin filaments cross-linked by fimbrin and its implications for bundle assembly and function. *J. Cell Biol.* **153**, 947–956.
48. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
49. Mukherjee, A. & Lutkenhaus, J. (1998). Purification, assembly, and localization of FtsZ. *Methods Enzymol.* **298**, 296–305.
50. Lockhart, A. & Kendrick-Jones, J. (1998). Interaction of the N-terminal domain of MukB with the bacterial tubulin homologue FtsZ. *FEBS Letters*, **430**, 278–282.
51. Oliva, M. A., Cordell, S. C. & Löwe, J. (2004). Structural insights into FtsZ protofilament formation. *Nature Struct. Mol. Biol.* **11**, 1243–1250.
52. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90–125, The Royal Society of Chemistry, Cambridge, UK.
53. Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575–588.

Edited by M. Moody

(Received 30 November 2006; received in revised form 7 March 2007; accepted 9 March 2007)

Circular and linear dichroism of proteins†‡

Benjamin M. Bulheller,[a] Alison Rodger[b] and Jonathan D. Hirst*[a]

Received 1st November 2006, Accepted 26th January 2007
First published as an Advance Article on the web
DOI: 10.1039/b615870f

Circular dichroism (CD) is an important technique in the structural characterisation of proteins, and especially for secondary structure determination. The CD of proteins can be calculated from first principles using the so-called matrix method, with an accuracy which is almost quantitative for helical proteins. Thus, for proteins of unknown structure, CD calculations and experimental data can be used in conjunction to aid structure analysis. Linear dichroism (LD) can be calculated using analogous methodology and has been used to establish the relative orientations of subunits in proteins and protein orientation in an environment such as a membrane. However, simple analysis of LD data is not possible, due to overlapping transitions. So coupling the calculations and experiment is an important strategy. In this paper, the use of LD for the determination of protein orientation and how these data can be interpreted with the aid of calculations, are discussed. We review methods for the calculation of CD spectra, focusing on semiempirical and *ab initio* parameter sets used in the matrix method. Lastly, a new web interface for online CD and LD calculation is presented.

1. Introduction

The theory of optical activity has come a long way since the phenomenon was first observed by Arago[1] in 1811. Jean-Baptiste Biot showed that the polarisation plane of light was altered after it passed through a quartz crystal.[2] Three years later, Biot reported similar rotation of the polarisation plane of linearly polarised light in several liquids, including turpentine and solutions of camphor.[3] Louis Pasteur, in 1848, interpreted these observations at the molecular level, at a time when molecules were not yet understood to be three dimensional.[4] Nevertheless, Pasteur showed that tartaric acid exists in two asymmetric forms, which rotate the polarisation plane of light in different directions.[5] In 1874, Le Bel[6] and van't Hoff[7] related rotatory power to the unsymmetrical arrangement of substituents at a saturated carbon atom, thus identifying the very foundation of stereochemistry. Through the definition of chirality, chemistry was given a mighty tool, which was able to explain the properties of sugars and many other organic compounds and which later led to the development of new analytical methodologies, such as optical rotatory dispersion (ORD)[8] and circular dichroism (CD) spectroscopy.[9]

Although the cause of optical activity was known, the development of a theoretical framework to describe and understand the phenomenon proved to be a complex task. The first adequate theory of optical rotatory power was presented by Born in 1915.[10] It was thoroughly investigated by Kuhn[11] and then reformulated by Rosenfeld in 1928,[12] who introduced the eponymous equation for the calculation of the rotational strength of a transition, which is related to its intensity in the CD spectrum.

Due to the sensitivity of CD and ORD to the secondary structure of proteins, the prediction of the optical spectra of polypeptides was attempted. Fitts and Kirkwood[13] calculated the optical rotation of a helical peptide in 1956 using polarisability theory, while Moffitt[14,15] in the 1950's used exciton theory.[16] Moffitt showed that the coupling of electric dipole allowed $\pi \rightarrow \pi^*$ electronic transitions in a helical arrangement leads to a resultant transition that is an in-phase combination, with a net polarisation parallel to the helix axis, and two transitions that are out-of-phase combinations, with a net polarisation perpendicular to the helix axis. Thus, he correctly predicted the right-handed nature of α-helices in proteins years before the first X-ray crystallographic structure of a protein. However, this approach was not readily developed into a quantitative method. In 1961, Doty established the dependence of the ORD on α-helical content and identified the electronic transitions of the peptide group as the most likely source of the rotatory power of proteins.[17] Later, he confirmed Moffitt's calculations[15] of the exciton splitting experimentally by resolving the three peptide electronic bands and attributing them to the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions, respectively.[18]

Building on the foundation laid by the aforementioned seminal studies, it has become feasible to calculate the CD spectra of molecules and today is, in fact, fairly routine for small molecules, for example, to determine the absolute configurations of isolated or synthesized compounds.[19–23] The computation of protein CD, however, remains a challenge, due to the size and flexibility of the molecules. For the calculation of the optical spectra of large molecules, such as

[a] *School of Chemistry, University of Nottingham, University Park, Nottingham, UK NG7 2RD. E-mail: jonathan.hirst@nottingham.ac.uk; Fax: +44 115 951 3562; Tel: +44 115 951 3478*
[b] *Department of Chemistry, University of Warwick, Coventry, UK CV4 7AL*

† Contract grant sponsor: EPSRC; Contract grant number: GR/T09224.
‡ The HTML version of this article has been additionally enhanced with colour images.

proteins and crystals, several methods have been developed specifically to cope with the size of the systems. The dipole interaction model[24–31] considers atoms and chromophores to act as point dipole oscillators, which interact through mutually induced dipole moments in the presence of an electric field. Another approach is the matrix method,[32] which we discuss in detail below. The following article reviews some of the developments in the calculation of optical properties of proteins, including both CD and the related polarised spectroscopy: linear dichroism (LD).

2. Circular dichroism

Among the related methods, ORD,[8] LD,[33] and CD,[9,34] the latter is now by far the most popular type of chiroptical or polarised light spectroscopy. The two energy regimes in current use are vibrational transitions[35–39] and electronic transitions. We focus only on the latter in this paper and refer to it by the abbreviation CD. When circularly polarised light impinges on a protein, the protein's electronic structure gives rise to characteristic bands in specific regions in the CD spectrum, reflecting the electronic excitation energies.[40,41] Secondary structural elements, such as α-helices, β-sheets, β-turns and random coil structures, all induce bands of distinctive shapes and magnitudes in the far-ultraviolet (Fig. 1).[42] For example, in an α-helix, an intense positive band at 190 nm and a negative band at 208 nm arise from the exciton splitting of electronic transitions from the amide non-bonding π orbital, π_{nb}, to the anti-bonding π orbital, π*, (Fig. 2). A negative band is located at about 220 nm, arising from the electronic transition from an oxygen lone pair orbital, n, to the π* orbital (Fig. 2). Other motifs give other spectroscopic shapes and signs.

The relative proportion of each secondary structure type can be determined by decomposing the far-UV spectrum into a sum of fractional multiples of the reference spectra.[9,43,44] For example, the average fractional helicity, f_H, of a peptide consisting of n residues can be determined from the observed mean residue molar ellipticity at 220 nm, $[\theta]_{220}$:

$$f_H = \frac{[\theta]_{220}}{[\theta_{H\infty}]_{220}\left(1-\frac{k}{N}\right)} \quad (1)$$

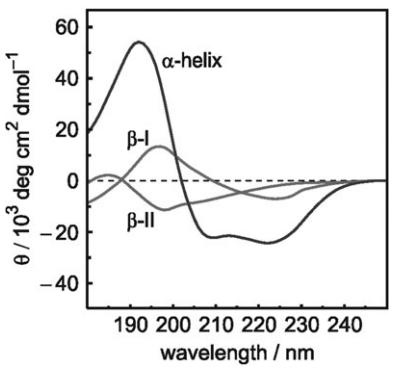


Fig. 1 Characteristic CD curves of secondary structure elements. The vertical axis shows intensity as the mean residue ellipticity, θ.

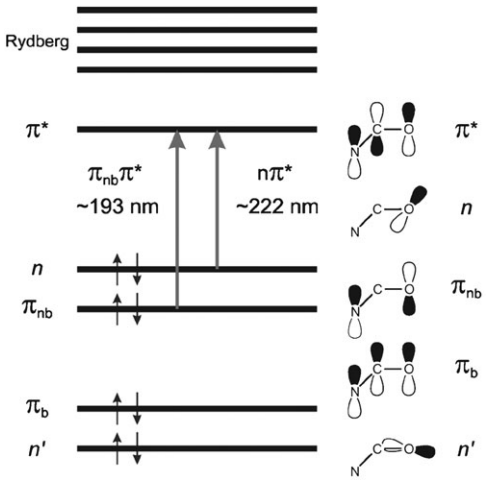


Fig. 2 Electronic transitions of the amide group in the far-UV region. The molecular orbitals shown are the bonding, nonbonding and antibonding π orbitals (π_b, π_{nb} and π*) and two lone pairs on the oxygen atom (n and n′).

where $[\theta_{H\infty}]_{220}$ is the ellipticity of a completely helical peptide of infinite length and k is an end-effect correction factor of approximately three.[45] Estimates of $[\theta_{H\infty}]_{220}$ range from −37 000 deg cm^2 dmol^{-1} [46] to −44 000 deg cm^2 dmol^{-1},[47] which corresponds to differential absorbance between −11.2 mol^{-1} dm^3 cm^{-1} and −13.3 mol^{-1} dm^3 cm^{-1}. However, $[\theta]_{220}$ can be influenced by several factors,[48–50] as will be discussed later. The empirical analysis of a spectrum can, therefore, lead to the determination of the protein secondary structure. Moreover, when CD spectroscopy is coupled with time-resolved experiments, protein folding events can be studied[51] and theoretical spectra can be used to interpret the results.[48,49] Although empirical fitting works remarkably well, due to the well-defined nature of the secondary structure motifs of proteins, it is important to proceed beyond empirical data analysis for a number of reasons. Since the conformation of proteins in solution may be (perhaps subtly) different from that in the crystalline environment, calculations can help to uncouple solution phase CD structure analysis from crystallographic structure motifs. Furthermore, the increasing importance of less well-folded or natively unfolded protein domains demands new approaches, since empirical comparison fails in these cases. In contrast to empirical data analysis, theoretical calculations, in principle, can also encompass the population dynamics of a solution of proteins and should be able to cope with new protein folds. The calculation of CD spectra from simulated ensembles of conformations can provide additional information from molecular dynamics (MD) simulations of protein folding.[52] The use of MD simulations, in particular, will be discussed later in more detail, but first we turn to the theory of CD and the associated computational methods.

2.1 Theory of CD

Circularly polarised light can be produced by the superposition of two linearly polarised light beams that are oscillating perpendicular to each other and propagating with a phase difference of π/2 radians. The magnitude of the electric field

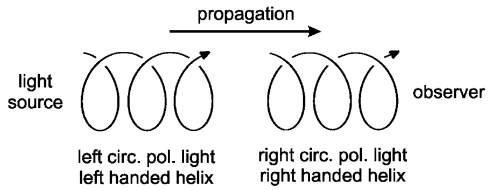


Fig. 3 Left and right, circularly polarised light, propagating in space.

vector of the resulting beam is constant, but rotates about the propagation direction. If the vector forms a right-handed helix, it is right circularly polarised light and *vice versa* (Fig. 3).[4,53] A stationary observer looking towards the light source at the electric field vector of right circularly polarised light sees it rotating in an anti-clockwise sense, if regarding its progression in space. However, more important is the dependence of the field vector with time, that is the direction of the field at the same position as a function of time. In this case, the vector of right circularly polarised light rotates clockwise (Fig. 4).[8,38]

A solution of chiral molecules possesses different refractive indices for left and right circularly polarised light, that is, the beams travel at different speeds and are absorbed to different extents at each energy. This means that the extinction coefficients for left and right circularly polarised light are different, $\varepsilon_L \neq \varepsilon_R$. This effect is called CD and the differential absorbance, $\Delta\varepsilon$, of left and right circularly polarised light (eqn (2)) is plotted against the wavelength λ to yield the CD spectrum.[9,43,54]

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R \tag{2}$$

The integral of $\Delta\varepsilon$ over a wavelength range associated with a particular transition is known as the CD strength or rotational strength of that transition. It is analogous to the oscillator strength of normal absorption. From the quantum electrodynamic viewpoint, the rotational strength, R^{0k}, of a transition from the ground state 0 to an electronically excited state k is the product of the electric transition dipole moment, $\vec{\mu}$, and the magnetic transition dipole moment, $\vec{m}$. The probability of a transition $0 \rightarrow k$ is represented by the integral $\langle\psi^k|\vec{\mu}|\psi^0\rangle$,

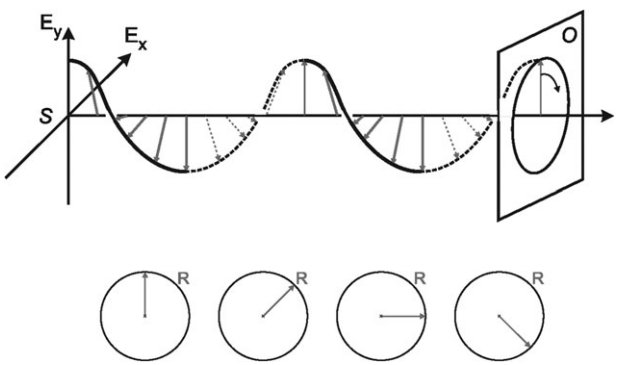


Fig. 4 Dependence of right circularly polarised light on the distance from the light source, *S*, (top) and on time (bottom) when the observer at position *O* is looking towards the source. The helix moves along the propagation direction without rotation, causing the vector of right circularly polarised light to rotate anticlockwise in space but clockwise in time.

which can be regarded as an oscillating dipole induced by the incident light beam.[43] $\vec{\mu}$ describes a linear displacement, whereas $\vec{m}$ characterizes a circulation of charge and accordingly the integral $\langle\psi^k|\vec{m}|\psi^0\rangle$ can be understood as a light-induced current loop. Thus, the combination of $\vec{\mu}$ and $\vec{m}$ creates a helical displacement of charge, leading to a different interaction with left and right circularly polarised light. While the operator for $\vec{\mu}$ is a real vector, that for $\vec{m}$ is imaginary, as it describes the rotation of charge in a complex coordinate system,[43] and it involves the linear momentum operator $\vec{p}$:

$$\vec{m} = \frac{e}{2mc}\left(\vec{r} \times \vec{p}\right), \quad \vec{p} = \frac{\hbar}{i}\nabla, \tag{3}$$

where e is the charge, m the mass of the electron, c the speed of light, $\vec{r}$ is the position operator of the electron, $\hbar$ is Planck's constant divided by 2π, $i = \sqrt{-1}$ and ∇ is the gradient operator. The rotational strength is given by the Rosenfeld equation:[12]

$$R^{0k} = \mathrm{Im}(\langle\psi^0|\vec{\mu}|\psi^k\rangle\langle\psi^k|\vec{m}|\psi^0\rangle), \tag{4}$$

where ψ^0 and ψ^k denote the wave functions of the ground state and the excited state, respectively, and Im the imaginary part of the product.

Using the Rosenfeld equation, it is possible to obtain the rotational strengths of a molecule and therefore calculate its CD spectrum. However, this requires wave functions for the ground and excited states, which due to computational limitations can only be determined *ab initio* for rather small compounds. Calculations on a protein, comprising hundreds of atoms and thousands of electrons, are thus a challenge, which demands some approximations. In the next section, we describe how these approximate wave functions may be calculated for proteins.

2.2 The matrix method

There are several methods to compute the CD spectra of proteins. In 1962, Tinoco adopted a perturbation approach, in which he considered the chromophores of a protein separately[55] and assumed that electrons are localized on a particular chromophore, with Coulombic forces the only means of interaction between different chromophores.[56] The matrix method,[32,57,58] which is derived from the Frenkel exciton model,[16] is commonly used for molecular crystals and chromophore aggregates. It is an improved formulation of the Tinoco[55] method and involves solving the eigenvalue problem *via* a matrix diagonalization rather than by applying perturbation theory and is hence more accurate (especially for degenerate and near-degenerate states) and easily implemented in computer algorithms. In the matrix method, orbitals on different chromophores are assumed not to overlap so that no interchromophore charge transfer occurs. The matrix method is fairly successful in calculating the CD spectra of proteins. Proteins with a high amount of α-helix can be calculated almost quantitatively.[59]

In the matrix method, the protein is split into M independent chromophores, with a monomer wave function ϕ_{is} for each chromophoric group i and electronic state s. The protein's kth excited state wavefunction, ψ^k, is then written (to a first approximation) as a linear combination of electronic

configurations, Φ_{ia}, in which only one chromophoric group i, is in an excited state a and the others are in the ground state, 0.[54] Thus

$$\Phi_{ia} = \phi_{10} \cdots \phi_{ia} \cdots \phi_{j0} \cdots \phi_{M0} \tag{5}$$

where ϕ_{ia} is the wave function of chromophore i after the transition $0 \rightarrow a$, and

$$\psi^k = \sum_{i}^{M} \sum_{a}^{n_i} c_{ia}^k \, \Phi_{ia} \tag{6}$$

The ground state of the protein is, similarly,

$$\psi^0 = \phi_{10} \cdots \phi_{i0} \cdots \phi_{j0} \cdots \phi_{M0} \tag{7}$$

Each transition in the protein CD spectrum is characterized by an energy and a rotational strength, which is given by eqn (4) and is related to the experimental intensity. To calculate the CD spectrum, we therefore need the wave functions ψ^k for each electronic excited state k of the protein (or at least those occurring in the spectral region of interest). The result of such a calculation is simply an intensity and a wavelength of each transition. To produce a spectrum we also need a bandshape and bandwidth for each transition. We return to this non-trivial issue below.

The first summation in eqn (6) is over the M chromophores; the second is over the n_i electronic excitations of each individual chromophore i, and c_{ia}^k are (expansion) coefficients, which have to be determined. To illustrate what this means in practice, consider a dipeptide with two excited states on each peptide, the lowest energy backbone n $\rightarrow$ π* and π $\rightarrow$ π* transitions. We then write

$$\psi^k = c_{1,\mathrm{n}\pi^*}^k \Phi_{1,\mathrm{n}\pi^*} + c_{1,\pi\,\pi^*}^k \Phi_{1,\pi\,\pi^*} + c_{2,\mathrm{n}\pi^*}^k \Phi_{2,\mathrm{n}\pi^*} + c_{2,\pi\,\pi^*}^k \Phi_{2,\pi\,\pi^*},$$

where

$$\begin{aligned} \Phi_{1,\mathrm{n}\pi^*} &= \phi_{1,\mathrm{n}\pi^*}\,\phi_{2,0} \\ \Phi_{1,\pi\,\pi^*} &= \phi_{1,\pi\,\pi^*}\,\phi_{2,0} \\ \Phi_{2,\mathrm{n}\pi^*} &= \phi_{1,0}\,\phi_{2,\mathrm{n}\pi^*} \\ \Phi_{2,\mathrm{n}\pi^*} &= \phi_{1,0}\,\phi_{2,\pi\,\pi^*} \end{aligned} \tag{8}$$

For each electronic excited state (k = 1–4) of the four-transition dipeptide, there is a different set of coefficients. The electronic excited states ψ^k of the protein and their corresponding energies can be calculated by solving the Schrödinger equation

$$\hat{H}\psi^k = E^k\psi^k \tag{9}$$

To solve eqn (9) we first need to construct the Hamiltonian, $\hat{H}$, which describes the dynamic properties of the system. If we denote the Hamiltonian of a local chromophore, i, as $\hat{H}_i$, then the Hamiltonian of the M independent chromophores is simply the sum of these local chromophore Hamiltonians. When the chromophores are allowed to interact, the Hamiltonian of the protein is the sum of all local Hamiltonians, $\hat{H}_i$, for each chromophore plus the sum of all intergroup potentials, $\hat{V}_{ij}$, of the entire molecule:

$$\hat{H} = \underbrace{\sum_{i=1}^{M} \hat{H}_i}_{\hat{H}_0} + \underbrace{\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \hat{V}_{ij}}_{\hat{V}} \tag{10}$$

Combining eqns (6), (9) and (10), one gets an equation in terms of chromophore wavefunctions and the coefficients $c_{i,a}^k$. It is convenient to convert this into a matrix formalism.

The matrix form of eqn (10) permits the calculation of the energy (the eigenvalues) and the wave functions (eigenvectors, that is, the coefficients $c_{i,a}^k$) by solving an eigenvalue problem through matrix diagonalization. The Hamiltonian matrix is diagonalised by a unitary transformation using the matrix U:

$$U^{-1}\hat{H}U = H_{\mathrm{diag}} \tag{11}$$

The diagonal elements of the resulting diagonal matrix (the eigenvalues) are the transition energies (excited state energies, since we set the ground state to be zero) of the interacting system and the eigenvectors, $c_{i,a}^k$, form the unitary matrix U. We then calculate the electric and magnetic transition dipole moments of the excitation from the ground state to the kth excited state as follows. The transition dipole moments of the non-interacting chromophores, which we denote $\vec{\mu}_a^0$ and $\vec{m}_a^0$, can be transformed to the interacting system using the unitary matrix from eqn (11):[54]

$$\begin{aligned} \vec{\mu}_i &= \sum_a U_{ai}\vec{\mu}_a^o \\ \vec{m}_i &= \sum_a U_{ai}\vec{m}_a^o \end{aligned} \tag{12}$$

From these transition dipole moments, the rotational strengths in the interacting system are readily calculated. For example, for a dipeptide and two transitions per group, the Hamiltonian matrix constructed using the Hamiltonian of eqn (10) has the form

$$\hat{H} = \begin{pmatrix} E_{1,\mathrm{n}\pi^*} & V_{1\mathrm{n}\pi^*;1\pi\,\pi^*} & V_{1\mathrm{n}\pi^*;2\mathrm{n}\pi^*} & V_{1\mathrm{n}\pi^*;2\pi\,\pi^*} \\ V_{1\mathrm{n}\pi^*;2\pi\,\pi^*} & E_{1\pi\,\pi^*} & V_{1\mathrm{n}\pi^*;2\pi\,\pi^*} & V_{1\pi\,\pi^*;2\pi\,\pi^*} \\ V_{2\mathrm{n}\pi^*;1\mathrm{n}\pi^*} & V_{2\mathrm{n}\pi^*;1\pi\,\pi^*} & E_{2\mathrm{n}\pi^*} & V_{2\mathrm{n}\pi^*;2\pi\,\pi^*} \\ V_{2\mathrm{n}\pi^*;1\pi\,\pi^*} & V_{2\pi\,\pi^*;1\pi\,\pi^*} & V_{2\mathrm{n}\pi^*;2\pi\,\pi^*} & E_{2\pi\,\pi^*} \end{pmatrix} \tag{13}$$

The above process is dependent on knowing values for the elements of the Hamiltonian matrix. This in turn requires wavefunctions and intergroup potentials for the independent chromophores. The diagonal elements of the matrix are the transition energies for each transition of each chromophore and the off-diagonal elements, V_{ij}, are the interaction energies between different transitions. These interactions are the cause for the dependency of protein CD spectra on secondary (and tertiary) structure. For the transitions $0 \rightarrow a$ on group i and $0 \rightarrow b$ on group j, the matrix element, V_{ij}, has the form:

$$V_{i0a;\,j0b} = \int_i \int_j \phi_{i0}\,\phi_{ia}\,\hat{V}_{ij}\,\phi_{j0}\,\phi_{jb}\,\mathrm{d}\tau_i\,\mathrm{d}\tau_j \tag{14}$$

If the interaction between the chromophores is regarded as an electrostatic interaction between charge densities ρ of separation r, that is $\hat{V}_{ij} = \frac{1}{4\pi\varepsilon_0 r_{ij}}$, then eqn (14) becomes:

$$V_{i0a;j0b} = \int_{r_i}\int_{r_j} \frac{\rho_{ia0}(r_i)\,\rho_{j0b}(r_j)}{4\pi\,\varepsilon_0\,r_{ij}}\,\mathrm{d}\tau_i\,\mathrm{d}\tau_j \qquad (15)$$

where $\rho_{ia0}(r_i)$ and $\rho_{j0b}(r_j)$ represent the transition electron densities on chromophores i and j, ε_0 is the vacuum permittivity and r_{ij} is the distance between the chromophores.

In principle, these matrix elements could be evaluated exactly from the monomer wave functions using the integral evaluation routines in many quantum chemical packages. In practice, an approximation is introduced to make the calculations tractable. In the *monopole–monopole approximation*, the permanent and transition densities are approximated by point charges and the integrals in eqn (15) are re-cast as a sum of the Coulomb interactions of these monopoles:[55]

$$V_{i0a;j0b} = \sum_{s=1}^{N_s}\sum_{t=1}^{N_t} \frac{q_s\,q_t}{r_{st}} \qquad (16)$$

where q_s and q_t are the point charges on chromophores i and j, and N_s and N_t are the number of these charges on the chromophore. Thus, at the core of matrix method calculations are the magnitudes and locations of the monopoles.[57,58] They reflect the orientation and magnitudes of the transition moments and they are critical for deducing the inter-chromophore interactions. Since the CD calculation depends only on these monopoles and their distance and orientation to each other, no *ad-hoc* definition of the secondary structure is needed. The amide group is the most important chromophore in protein far-UV CD and has therefore been a key target for research into protein CD.

2.3 Calculation of CD

The implementation of the matrix method as a computer program is quite straightforward. The computation requires the coordinates of the atoms and the positions of the chromophores in the protein. Each chromophore is parameterised by a set of monopoles describing the electrostatic potential and for every group in the protein the respective set of monopoles is superposed on the chromophore's atoms. By calculating the interaction between all the different electronic excitations, the Hamiltonian matrix is constructed and, as described above, the rotational strengths are determined, yielding a line spectrum, that is, values for the rotational strength of each transition. However, in an experimental spectrum, the transitions are broadened due to the uncertainty principle, unresolved vibronic components and the interaction of the chromophore with its environment including other chromophores and the solvent. Thus, an overlay of approximately Gaussian shaped bands is observed. Since the result of the calculation is a line spectrum, a convolution with a lineshape function is required. Gaussian line shapes give better results than Lorentzian curves. Hence, all the spectra shown in this paper have been created by taking a Gaussian function of the type[40]

$$y = y_0 e^{-U_0^2} \quad \text{with} \quad U_0 = \frac{\lambda - \lambda_0}{\Delta_0} \qquad (17)$$

The band width or full width at half maximum height (fwhm) of this function is set to a single fixed value of 12.5 nm for all transitions in this work. The resulting calculated spectrum can then be compared to the experimental spectrum.

2.4 Comparison of different parametrisations of the amide chromophore

As discussed above, for expedience the Hamiltonian matrix elements may be estimated using the monopole–monopole approximation, which involves representing the transition densities of a chromophore by a set of point charges. Two modern parameter sets for matrix method calculations are an *ab initio* derived set, NMAabinit,[60] and the set of Woody and Sreerama,[61] NMAsemi. They arise from different approaches and are discussed in detail below.

The NMAabinit set has been derived entirely from *ab inito* calculations. Using *ab initio* techniques, the amide chromophore was parametrised with N-methylacetamide (NMA) as a model compound.[60] The electronic spectrum of NMA in solution was calculated, using the complete-active space self-consistent-field method implemented within a self-consistent reaction field (CASSCF/SCRF),[62–65] combined with multi-configurational second-order perturbation theory (CASPT2-RF).[62,65] Monopoles for a given state were determined by fitting their electrostatic potential to reproduce the *ab initio* electrostatic potential for that state so that the least-squares difference is minimised (typically within 5%). The parameter set[60] consists of 32 monopoles for the amide n → π* transition all at a distance of 0.1 Å from the C, N, O and H atoms, as shown in Fig. 5, and 20 monopoles for the π → π* transition, with one charge situated at each atom centre and four around each atom centre at a distance of 0.05 Å (Fig. 5). Two higher energy transitions at 129 and 123 nm were also parametrised.

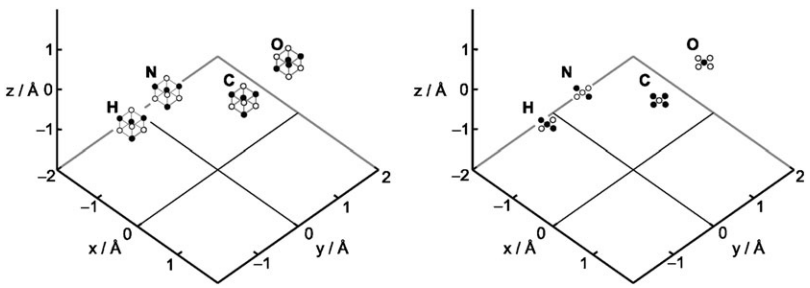


Fig. 5 Monopole positions of the *ab initio* amide parameter set, NMAabinit, for the n → π* transition (left panel) and the π → π* transition (right panel). Negative charges are represented by black circles and positive charges by white circles. The distance of the monopoles to the atoms is about 0.09 Å in the left panel and either 0 or 0.05 Å in the right panel.

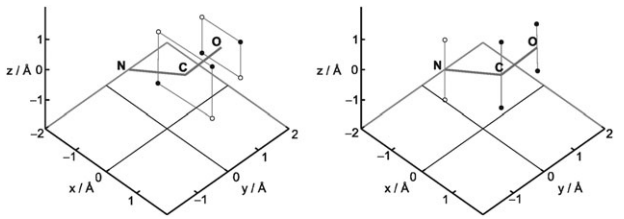


Fig. 6 Monopole positions of the semiempirical amide parameter set, NMAsemi, for the n → π* transition (left panel) and the π → π* transition (right panel). Negative charges are represented by black circles and positive charges by white circles. The distance of the monopoles to the atoms is between 0.8 and 1.2 Å.

NMAabinit, has been tested on a set of 47 proteins,[66] and the calculated CD agreed well with experiment, particularly at 220 nm, where the accuracy was nearly quantitative. This wavelength is important, since the intensity of the band in that region correlates with the helical content of the protein.[67]

NMAsemi is derived from a combination of semiempirical calculations using the INDO/S method and experimental data. The n → π* transition density is represented by four monopoles surrounding each of the carbon and oxygen atoms, whereas for the π → π* transition density a total of six charges are located above and below the peptide plane (Fig. 6). Compared to the *ab initio* set, it consists of fewer monopoles with smaller charges at greater distance from the atomic centres. However, the electrostatic potential created by the charges is the physically relevant quantity, rather than the particular values and locations of the charges, which are (within certain constraints) arbitrary, as an infinite number of monopole sets give rise to the same electrostatic potential.

For the peptide chromophore, the electrostatic potentials for the two parameter sets are compared in Fig. 7, which shows a contour plot of the electrostatic potential arising from the n → π* transition density 0.5 Å above the plane of the peptide group. In NMAsemi, the monopoles located around the carbon and oxygen produce a symmetric potential around the CO bond. Due to the additional monopoles at the nitrogen and hydrogen atoms, the electrostatic field generated by NMAabinit asymmetric. The most significant difference between the two parameter sets lies indeed with the n → π* transition, as the π → π* transitions are highly dipolar and look very similar in shape and magnitude. However, there is a difference in the orientation of the π → π* transition dipole moment of 8.5°, and the importance of this orientation has been discussed in detail in the literature.[59,61]

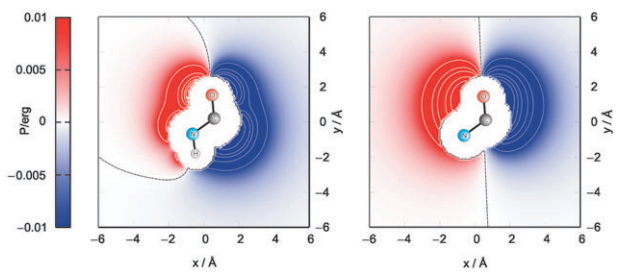


Fig. 7 Contour plots of the electrostatic potential *P* of the n → π* transition caused by a peptide chromophore for the *ab initio* parameter set, NMAabinit (left panel) and for the semiempirical parameter set, NMAsemi (right panel). Solid lines denote a positive, dotted lines a negative potential. The contour line of the zero-crossing is dashed.

2.5 Influence of the side chains

In this paper, we concentrate on CD calculations, which consider only the peptide chromophore. However, there are other chromophores of interest. The chromophoric groups of the side chains dominate the CD spectrum in the near-UV (250–350 nm). The intensity in this region is about two orders of magnitude lower than in the far-UV in most proteins, reflecting the fact that only a few percent of the residues in any protein contain aromatic groups. For the same reason, the backbone amide transitions dominate the far-UV in most proteins and we only observe the aromatic groups clearly above 250 nm, where the amide group has no electronic transitions. In the near-UV, as in the far-UV, chromophore parameter sets based on *ab initio* or semiempirical calculations on individual chromophores may be constructed. Our approach has been to gain a more accurate description of the valence electronic transitions of the side chain chromophores by performing *ab initio* calculations.[68] Benzene, phenol and indole are the chromophoric groups of the amino acids phenylalanine, tyrosine and tryptophan, respectively. The influence of solvent on the most important transitions of each chromophore has also been investigated.[68] The excited states are by convention labelled as ^1L_b, ^1L_a, ^1B_b and ^1B_a, in which the superscript indicates that the state is a singlet state and the subscript indicates the orientation of the electric transition dipole moment as along either (a) the long or (b) the short axis of the ring in phenylalanine and tyrosine.[69] L refers to the low-lying transitions and B to the transitions at high energy, respectively. In tryptophan the moments are inclined to the long axis of the indole chromophore at about +54° (^1L_b), −41° (^1L_a), +18° (^1B_b) and −61° (^1B_a), respectively.[70] The four states of phenylalanine, tyrosine and tryptophan are responsible for the majority of the spectroscopic bands due to side chains in proteins. The ^1B states are more intense than the ^1L states and are located in the far-UV region. *Ab initio* calculations reproduced the solvatochromatic shifts of the individual chromophores well, although the oscillator strengths were underestimated for benzene and vibronic coupling was neglected. The derived parameter set was tested on the near-UV spectra of 30 proteins and on the difference spectra of 20 mutants and in both cases gave a significant improvement on the semiempirically derived parameters.[71]

2.6 Comparing theory and experiment

Fig. 8 (left panel) compares calculated spectra using the NMAabinit and NMAsemi parameters with the experimental spectrum of a helical protein. The peak around 193 nm is well matched by the *ab initio* parameter set, whereas the peak of NMAsemi is slightly shifted to higher energy. In the region between 200 and 230 nm NMAsemi indicates the double minimum which is characteristic for a helical protein, but which is not resolved for NMAabinit. For the β-sheet protein

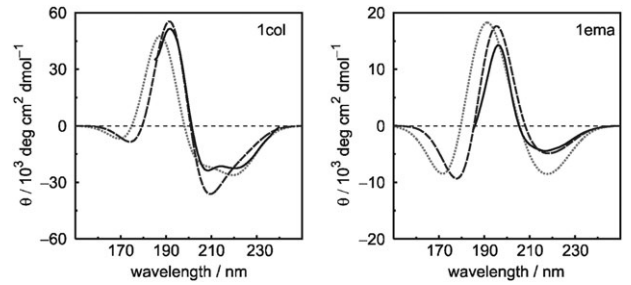


Fig. 8 Comparison of the experimental spectra of colicin A (1col, mostly α-helical, left panel) and green fluorescent protein (1ema, β-I class, right panel) with the spectra calculated using two different parameter sets. Experimental (solid), *ab initio* parameter set (dashed), semiempirical parameter set (dotted).

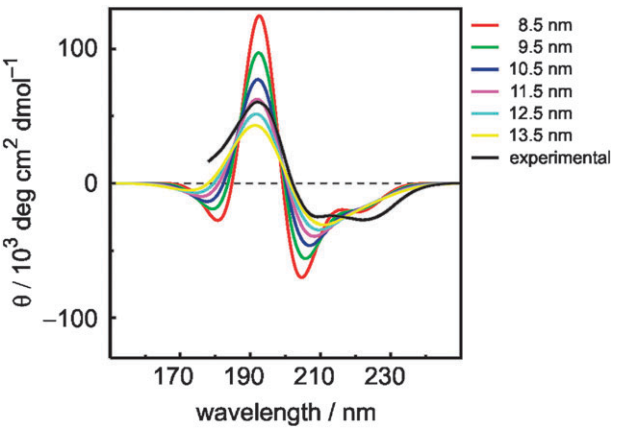


Fig. 9 Calculated CD spectra where all transitions have been assigned the same bandwidth, with values ranging from 8.5 to 13.5 nm.

(Fig. 8, right panel) the intensity at 195 nm is overestimated by both parameter sets, whereas NMAabinit shows a better fit at longer wavelengths, including again a good estimation at 220 nm. In β-II proteins,[72] the strands are rather short or not aligned in a parallel manner, but often twisted and bent. The resulting spectra are similar to unfolded, random coil proteins (Fig. 1) and the agreement of the calculations with experiment is poor. Around 220 nm the intensity is reproduced well, but at higher energies, especially at 200 nm where the minimum is not predicted at all, no correlation between theory and experiment is achieved.

Another structure occurring in peptides is the poly(proline)II-helix (P_{II}, $\phi = -75°$, $\psi = +145°$). The P_{II} conformation is the subject of much current research and has been identified as a possible origin for the distinction between β-I and β-II in β-rich proteins. Sreerama and Woody[72] related the observation of different types of spectra for β-I and β-II proteins back to the content of P_{II}-conformations, *i.e.* proteins possessing a ratio of P_{II} to β-sheet less than 0.4 appear to belong to the β-II-type, whereas β-II-proteins show a ratio of P_{II} to β-sheet greater than 0.4.[72] The CD spectrum of the P_{II} conformation exhibits an intense negative band at 200 nm, which is not reproduced in calculations.[54] Another influential structural factor is the number of disulfide bridges. The agreement between calculated and experimental spectra deteriorates with more irregular structure and greater numbers of disulfide bonds.

In order to enhance the quality of the calculations of CD spectra, many aspects have been studied, for example, the inclusion of additional chromophores[68] and investigations into hydrogen bonding.[73] Another parameter which has been examined is the width of each band at a specific transition energy. In most calculations, this is set to a single fixed value for all wavelengths. Wavelength-dependent bandwidths have also been tested, but can lead to unexpected results. A comparison of the resulting CD spectra when all transitions were assigned the same bandwidth between 8.5 and 13.5 nm are compared in Fig. 9. To improve resolution of the two peaks around 208 and 220 nm for helical proteins, bandwidths between 7.5 and 15.5 nm have been explored.[66] The minimum of the n → π* transition at 220 nm is resolved for values below 12.5 nm, but such narrow bandwidths then exaggerate the intensity maximum of the peak at 190 nm. In reality, different transitions may well have different bandwidths, but, without any justification for a particular choice, there is a danger of introducing a number of empirical parameters which artificially enhance agreement between theory and experiment. In the absence of a rigorous reason for choosing different bandwidths for different transitions, we adopt the best overall value (12.5 nm) in all of our calculations, while recognizing that we are thereby almost certainly worsening the agreement of calculation and experiment.

To assess the two different parameter sets, the Spearman rank correlation[74] between the experimental and calculated intensity has been determined at each wavelength for a set of 46 proteins (Fig. 10). α-Bungarotoxin (2abx) has been removed from the initial set, because an analysis using PROCHECK[75] showed that only 14.8% of the residues lie in the core regions of the Ramachandran plot and 24.6% in disallowed regions, thus questioning the structural integrity of the file. At 190 nm the NMAsemi parameters give a correlation of 0.68, compared to 0.85 of NMAabinit. In this region, the

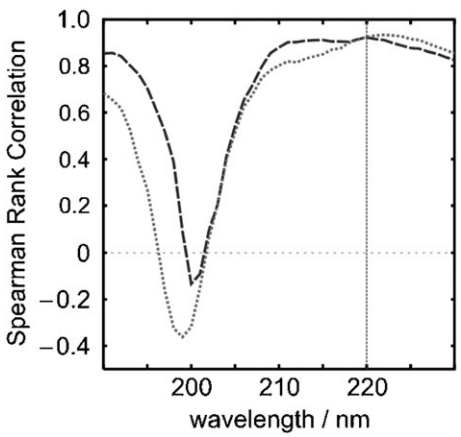


Fig. 10 Spearman rank correlation between the calculated and experimental intensity at a particular wavelength for a set of 46 proteins for the *ab initio* parameter set (NMAabinit, dashed) and the semiempirical parameter set (NMAsemi, dotted).

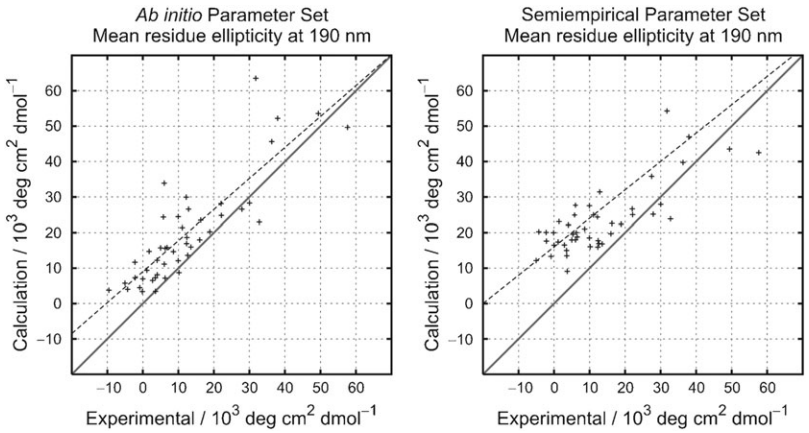


Fig. 11 Comparison of experimental and calculated mean residue ellipticity at 190 nm. The solid line corresponds to perfect correlation. The dashed line shows the linear regression through all points.

intensity is mostly overestimated by both sets (Fig. 11), but the overall correlation is positive. Notably, bacteriorhodopsin (2brd) is an outlier, despite being mostly α-helical. It is a membrane protein with a high helical content. Its CD spectrum is less intense than the spectra of globular proteins with similar helical content and, thus, the intensity of the spectra is overestimated by both parameter sets.

The point where the CD intensity changes sign, at around 200 nm, is a challenge for both parameter sets. In this region the gradient of the spectrum is greatest, causing a high difference in intensity for even small deviations of the calculated *versus* the experimental wavelength. Although the absolute error in the wavelength where the intensity is calculated to be zero is below 3 nm (Table 1), the intensities may differ up to 10 000 deg cm^2 dmol^{-1}. Hence, the correlation is almost zero for NMAabinit and is actually negative for NMAsemi. NMAabinit tends to overestimate the intensity, *i.e.* the calculated wavelength of the zero point is too large and the spectrum slightly red-shifted, whereas the deviations of NMAsemi are equally distributed. The correlation steadily increases after the zero-crossing and both parameter sets reach almost constant values of between 0.8 and 0.9 above 210 nm. The band at 208 nm, arising from the exciton splitting of the $\pi \rightarrow \pi^*$ transition, shows a correlation of 0.80 and 0.75 for NMAabinit and NMAsemi, respectively. At 220 nm, which is important for the determination of helical content, both sets give a comparably (high) correlation (Fig. 12).

2.7 Applications and other developments

In protein folding the transient and short-lived nature of intermediates makes the direct observation of folding and unfolding exceptionally challenging for techniques such as NMR and X-ray crystallography. However, nanosecond CD spectroscopy is a promising candidate for monitoring the conformational changes, which motivates the combination of these first-principles calculations with MD simulations to study the influence of dynamics on the CD of polypeptides.[52] MD simulations and CD calculations have been performed on concanavalin A, a β-I- and elastase, a β-II-type protein, which possess similar structures but show different CD spectra.[52] In the former case, the spectrum did not change substantially and remained close to that calculated for the X-ray crystal structure. For the β-II protein, the intensity of the calculated positive peak at 195 nm (which is negative in the experiment) decreased considerably, suggesting that relaxation of the protein structure in solution may be a contributing factor to the differences between β-I and β-II proteins. The well ordered, aligned sheets in β-I proteins and also the helices in α-helical proteins induce a relatively rigid structure supported by main-chain hydrogen bonds. This framework cannot be formed to the same extend in less ordered peptides such as β-II-type proteins, making them more flexible.

The effect of hydrogen bonding on CD is of interest, and this was highlighted by the unexpectedly high negative intensities at 220 nm reported for several peptides, which exceed even the highest estimate of $[\theta_{H\infty}]_{220}$, the intensity expected for a 100% infinitely long helix.[45] Calculations suggested that the interaction with highly polar solvent molecules, which bind selectively to the helical conformation of peptides, may strengthen the backbone hydrogen bonds of the peptide. The intensity at 220 nm is sensitive to the main-chain hydrogen bond length[49] and the stabilization (and shortening) of hydrogen bonds could thus cause an over-estimation of the helicity in the presence of polar solvents.

Peptides are often studied as model systems for probing interactions that are important in protein stability and folding. Many aspects of the stability of α-helices have been investigated, including the interactions between aromatic and basic side chains.[48] Often a tyrosine residue is introduced into the peptide to facilitate concentration measurements. Calculations of protein CD from first principles indicate that the presence of tyrosine can have a direct effect on the CD and can change the intensity at 222 nm up to ±5000 deg cm^2 dmol^{-1}, depending on the orientation of the aromatic ring.[48] Further investigation into the influence of tyrosine has shown that the intensity at 220 nm for tyrosine-containing peptides will lead to underestimates of the helical content of the peptide of between 5 and 20%.[50]

The deep-UV is a region of the spectrum of growing interest for CD. The use of synchrotron radiation has made it possible

Table 1 Calculated (using NMAabinit parameters) and experimental wavelengths of interest in protein CD. The 13 proteins at the top show the least agreement between the calculated and experimental spectra; β-II proteins are printed in bold

Protein name	Protein PDB code	λ_{max}/nm		$\lambda_{[\Delta\varepsilon=0]}$/nm		λ_{min}/nm	
		Calc	Exp	Calc	Exp	Calc	Exp
Carbonic anhydrase	**1ca2**	195	—	207	—	218	—
Bence-Jones immunoglobulin	1rei	199	202	215	210	223	217
Bacteriorhodopsin	2brd	191	196	201	205	209	—
Chymotrypsinogen A	**2cga**	196	—	209	—	219	—
Ribonuclease S	2rns	195	—	206	—	217	—
Cu,Zn Superoxide dismutase	**2sod**	199	190	212	196	222	209
Erabutoxin B	3ebx	199	201	217	209	223	215
Native elastase	**3est**	197	—	211	—	221	—
Trypsin	**3ptn**	198	—	212	—	222	—
Gamma-B Crystallin	4gcr	197	203	211	—	221	—
Alpha Chymotrypsin A	**5cha**	197	—	211	—	220	—
Trypsin inhibitor	**5pti**	193	—	203	—	211	—
Beta-Lactoglobulin	1beb	194	190	205	201	216	213
Colicin A	1col	192	192	201	201	210	208
Azurin	1e5z	196	195	209	205	219	219
Green fluorescent protein	1ema	195	195	208	206	218	217
Eco Ri Endonuclease	1eri	192	190	202	201	211	208
Flavodoxin	1fx1	193	194	204	203	214	214
Intestinal fatty acid binding protein	1ifc	195	197	207	204	217	216
Lysozyme	1lys	192	192	202	199	211	208
Plastocyanin	1plc	198	200	213	—	222	—
Rhodanese	1rhd	194	193	204	201	213	210
Subtilisin	1sbt	192	192	202	200	212	218
Cachectin	1tnf	197	201	211	209	221	221
Concanavalin A (native)	2ctv	198	197	212	209	221	224
Hemerythrin (adizomet)	2hmz	192	194	201	202	210	210
Hemoglobin (horse)	2mhb	191	195	201	202	210	210
Prealbumin (human plasma)	2pab	197	195	210	202	221	217
Pepsinogen	2psg	195	193	207	200	218	211
Staphylococcal nuclease	2sns	193	–	203	199	211	209
Adenylate kinase	3adk	193	192	203	202	211	210
D-Glyceraldehyde-3-phosphate dehydrogenase	3gpd	192	192	201	200	211	212
Glutathione reductase	3grs	194	191	204	199	212	207
Phosphoglycerate kinase	3pgk	192	191	202	200	211	209
Porin	3por	194	195	205	208	214	218
Ribonuclease A	3rn3	194	193	205	199	215	209
Triosephosphate isomerase	3tim	192	193	202	202	211	219
Insulin	4ins	190	195	199	202	208	209
Myoglobin	4mbn	192	192	201	201	210	209
Apo-Liver alcohol dehydrogenase	5adh	194	196	205	203	217	211
Carboxypeptidase	5cpa	193	198	203	204	212	211
Calcium-binding parvalbumin B	5cpv	191	195	200	202	209	209
Cytochrome C	5cyt	194	195	205	202	214	210
Apo-Lactate dehydrogenase	6ldh	193	192	203	202	212	212
Thermolysin	8tln	193	194	203	203	211	211
Papain	9pap	192	188	201	194	211	209

to extend the wavelength range over which data can be measured into the vacuum-UV[76] down to 160 nm to uncover further features of the protein CD spectra.[77–82] The origin of a band around 165 nm has prompted some speculation. At first a transition from the higher-energy lone-pair orbital on oxygen to an antibonding σ orbital (n → σ*) was suggested.[83] This was discounted by *ab initio* calculations, which showed that the σ* orbital lies at a much higher energy.[84] The involvement of an excitation from the bonding π_b orbital to π* (π_b → π*) was suggested,[85] but this transition occurs around 140 nm.[84] An excitation from the second lone pair on oxygen to the anti-bonding π orbital (n′ → π*) has also been considered. Alternatively, there is experimental and theoretical evidence indicating that charge-transfer transitions between neighbouring amide groups are the cause of the band at 165 nm.[86–88]

To investigate this, we have performed CASSCF/CASPT2 calculations on 13 peptide dimers[41,82,89] with dihedral angles from all accessible areas of the Ramachandran plot. The charge-transfer between neighbouring groups had a considerable impact on the calculated spectrum in a manner dependent on the conformation, that is the dihedral angles of the two amide groups. For a matrix method calculation this means that each conformation has to be parametrised. The Hamiltonian matrix needs to be extended by the addition of a charge-transfer chromophore, which spans two neighbouring monomers in the chain.

Another area of investigation is the inclusion of higher energy transitions, since interaction with these could affect the intensities and positions of the bands in the far-UV. The π_b → π* (bonding π orbital to π*) and n′ → π* (second lone pair on oxygen to π*) transitions occur around 140 nm and are thus

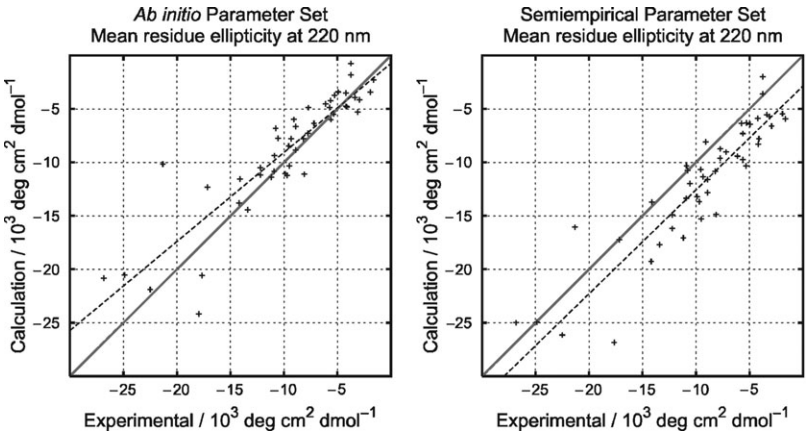


Fig. 12 Comparison of experimental and calculated mean residue ellipticity at 220 nm. The solid line corresponds to perfect correlation. The dashed line shows the linear regression through all points.

outside the measured region of 180–250 nm. Nevertheless, they could couple with the n → π* and π → π* transitions and have an effect on the observed far-UV. However, including the two higher energy transitions decreases the intensity of the peak at 193 nm and shifts it to longer wavelengths. The calculations show a negative peak at 182 nm (which is not apparent in experimental spectra) and the inclusion of the two additional transitions even increased its intensity. The bands around 208 nm and 220 nm are practically unaffected and the correlation did not improve in this region. Thus, the consideration of the two higher energy transitions did not lead to an improvement of the agreement.[66]

3. Linear dichroism

3.1 Introduction to LD

Linear dichroism (LD) is also a differential absorption technique, but in this case it is the difference in absorption of light linearly polarised parallel and perpendicular to an *orientation* axis (Fig. 13):[90]

$$\mathrm{LD} = A_{\parallel} - A_{\perp} \tag{18}$$

LD is related to CD in that both require the difference between the absorbances of different polarised light beams to be measured and CD spectropolarimeters can be adapted to produce the required alternating beams of linearly polarised light for LD. However, LD measurements are performed on systems that are either intrinsically oriented or are oriented during the experiment.

In order to understand LD, one has to consider what happens when a molecule absorbs a photon: absorption can be pictorially viewed as either the electric field or the magnetic field (or both) of the radiation pushing the initial electron density to a final arrangement of higher energy. The direction of net linear displacement of charge is known as the *polarisation of the transition*. The polarisation and intensity of a transition are dependent on the electric transition dipole moment $\langle \psi^0 | \vec{\mu} | \psi^k \rangle$, which may be regarded as the antenna by which the molecule absorbs light. Each transition thus has its own antenna and the maximum probability of absorbing light is obtained when the antenna and the electric field of the light are parallel. Now imagine we have a linearly polarised light beam (Fig. 13) and a sample of molecules all oriented in exactly the same way. If we first measure the normal absorption with the light polarised so that it is parallel to the direction of orientation of the sample and then measure it when the light is polarised perpendicular to this direction, the difference of these two spectra is the LD spectrum. The two extreme situations we might encounter are:

(1) If the polarisation of the transition we are probing is perfectly *parallel* to the orientation direction, then

$$\mathrm{LD} = A_{\parallel} - A_{\perp} = A_{\parallel} > 0 \tag{19}$$

(2) If the polarisation of the transition we are probing is *perpendicular* to the orientation direction, then

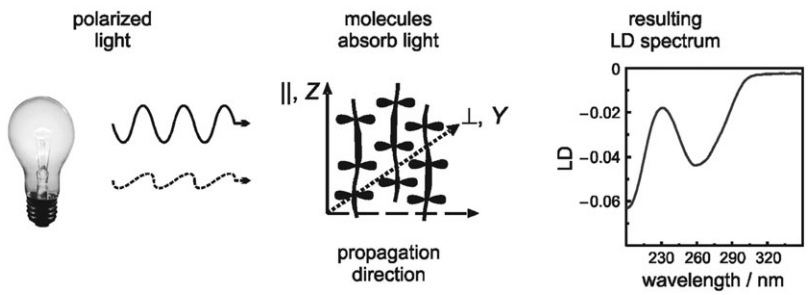


Fig. 13 Schematic illustration of an LD experiment.

$$\mathrm{LD} = A_{\parallel} - A_{\perp} = -A_{\perp} < 0 \qquad (20)$$

For intermediate polarisations, the LD is between these cases and for a perfectly oriented system:[90]

$$\mathrm{LD}^r = \frac{\mathrm{LD}}{A} = \frac{A_{\parallel} - A_{\perp}}{A} = \frac{3}{2}(3\cos^2\alpha - 1) \qquad (21)$$

where the reduced LD, LD^r, is defined to be the ratio of the LD over the isotropic absorbance, A. (LD^r is independent of sample pathlength and concentration) and α is the angle between the transition moment and the orientation axis. Thus, we can determine the polarisation of a given transition from its LD spectrum if we know how the molecule is oriented; conversely, we can use LD as a probe of molecular orientation if we know the polarisation of a transition moment within the molecule.

3.2 Molecular alignment techniques

In addition to the requirement of linearly polarised light for LD, LD experiments are technically demanding because of the need to orient the sample. In order to appreciate how we can use LD to study proteins, it is important to understand what data can be collected. This depends on how samples can be aligned. Proteins may be oriented by a number of methods depending on their nature. Long polymers, such as fibrous proteins, may be oriented by the viscous drag caused when a solution is flowed between narrow walls.[91–93] Depending on the cell design, the light is then propagated either along the flow direction or perpendicular to it. This technique is commonly used for LD studies of DNA. In the 1960's and 70's it was used for a few fibrous protein studies.[94,95] More recently we have been developing it for a wide range of fibrous proteins including actin, tubulin, FtsZ, collagen and amyloid fibres.[96–99] The most successful flow cell has proved to be a cylindrical Couette flow cell, where the solution containing the long molecules is subjected to a constant gradient over the annular gap between two coaxial cylinders one of which is rotating. The advantage of the Couette cell over a flow through system is that the same volume of sample can be measured indefinitely. For proteins a significant step forward has been the recent development of micro-volume cells which only require 30 μL of sample.[100]

Couette flow orientation of proteins has been extended to membrane proteins and peptides following Ardhammar, Mikati and Nordén discovering[101] that spherical liposomes were distorted in a Couette cell and could be used to orient small molecules bound to the liposomes. We later showed that any protein or peptide that bound to a liposome could be similarly oriented.[33,102] Due to the defined symmetry of the liposome, for such a system eqn (21) is replaced by:

$$\mathrm{LD}^r = \frac{3S}{4}(1 - 3\cos^2\beta) \qquad (22)$$

where β is the angle the transition moment of interest makes with the normal to the liposome surface (*i.e.* parallel to the *lipid* long axis) and S is the orientation factor that denotes the fraction of the liposome that is oriented as a cylinder perfectly parallel to the flow orientation.

3.3 Examples of protein LD data

Fig. 14 shows the LD spectrum one might expect to measure in the absence of light scattering artefacts for an α-helix oriented parallel to the sample orientation direction. This might, for example, be on the surface of a liposome or a coiled-coil α-helical fibre. The 220 nm region is dominated by transitions polarised perpendicular to the helix and therefore shows a negative signal; the 208 nm region is polarised parallel to the helix and so it has a positive signal. However, it is sandwiched between two negative signals so it may or may not actually end up looking like a positive maximum as the signals overlay and cancel. This is an illustration of why we need to have electronic level calculations of LD spectra, as discussed below, in order to deconvolute spectra in structural terms.

Fig. 15 shows a near- and far-UV LD spectrum of actin over a range of concentrations.[93] In the 12 μM spectrum one can see the components of the $\pi \rightarrow \pi^*$ transition and the $n \rightarrow \pi^*$ transition of the peptide backbone, comparable to the schematic of Fig. 14, but of opposite sign. The lower energy (higher wavelength) component of the $\pi \rightarrow \pi^*$ band is evident as the 215 nm dip in the spectrum. The LD spectrum therefore indicates that, on average, the α-helices in the actin fibre are oriented more perpendicular than parallel to the fibre axis. This is in accord with literature data for actin. However, if we did not have literature data it is almost certain we would want a less qualitative description of the structure, which is simply not possible without considering all the transitions in the protein as discussed below.

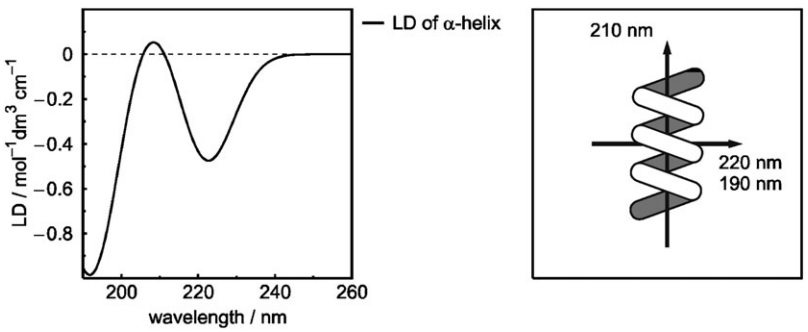


Fig. 14 A schematic illustration of the LD spectrum one might expect for an α-helix oriented parallel to the sample orientation direction. α-helices have an $n \rightarrow \pi^*$ transition polarised perpendicular to the helix axis at 220 nm;[33] a component of the $\pi \rightarrow \pi^*$ transition parallel to the helix at ~210 nm and a component of the $\pi \rightarrow \pi^*$ transition perpendicular the helix at ~190 nm.[33,97]

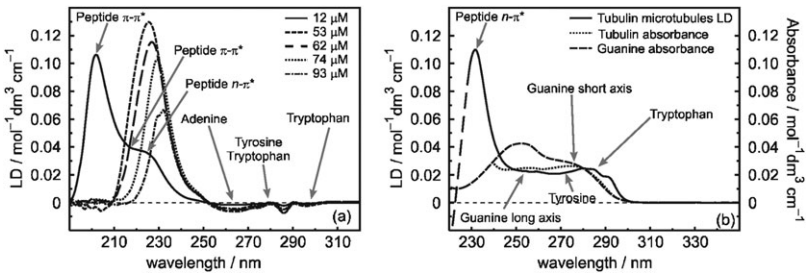


Fig. 15 (a) Near and far-UV LD spectra of F-actin in a Couette flow cell showing the apparent shift to shorter wavelength of the maximum signal as the concentration of F-actin is reduced. The 12 μM spectrum (solid line) is 'true'. (b) Near-UV absorbance and LD data for tubulin (microtubules include tubulin and guanine).

It should be noted that care must be taken in collecting low wavelength data for a system such as actin where light scattering is significant, as illustrated in Fig. 15. In that case more concentrated samples have an apparent wavelength shift of the maximum LD signal of ~30 nm to longer wavelengths. When the dilution effect is simply to reduce the signal intensity according to the Beer Lambert law, then we can conclude the spectrum is real and one might expect a calculation to be able to give the same result if the geometry is correct. Matrix method calculations (see below) will never reproduce artefacts in spectra.

Fig. 15b shows the near-UV LD and absorbance spectra of tubulin.[99] Due to the different wavelengths of guanine, tryptophan and tyrosine transitions, comparisons of calculated spectra with experimental data should enable one to determine, for example, the orientation of the guanine (which is essential for polymerisation) at all stages of tubulin polymerisation/depolymerisation. Thus, calculations on tubulin (and indeed the near-UV region of actin) will require accurate parametrisation of the NTP/NDP involved in the polymerisation as well as the protein chromophores. This is true for any additional chromophore in any protein, both in CD (as discussed above) and LD.

In the case of bacteriorhodopsin (BR), the retinal chromophore is a very useful probe of extent of protein orientation. The absorption, CD and LD spectra of BR in liposomes are given in Fig. 16.[103] The absorption spectra (Fig. 16a) are consistent with those reported in the literature for the initial state of BR.[104] The absorption maximum at ~570 nm is due to a long axis polarised transition of the retinal chromophore. The broad peak in the near-UV region (260–290 nm) is due to the transitions of the protein aromatic side chains phenylalanine, tyrosine and tryptophan and the peak observed in the far-UV region (215 –~230 nm) is due to the peptide n → π* transition of the amide groups. The backbone CD spectrum of BR (Fig. 16b) is in accord with the CD spectrum for a highly α-helical protein, although the intensity is weaker than for globular proteins as discussed above.[105,106] The ~570 nm LD signal is consistent with a retinal chromophore oriented at ~70° to the lipids of a lipid bilayer if S ~0.05. The aromatic region (260–280 nm) of the LD spectrum is dominated by the indole chromophore of the tryptophan residues.[107] Contributions from L_a (270 nm) and L_b (287 nm) transitions are apparent, both showing positive LD and suggesting the average tryptophan is tilted so that the normal to the plane of the indole is ~40° from the average lipid and the long and short axes are both at an angle of 60–65° to the lipids. This is consistent with the fact that the retinal is sandwiched by tryptophan residues in the X-ray structure.[105]

The protein backbone LD spectrum shows a positive maximum at 220 nm (n → π*) and a negative maximum at ~213 nm (π → π*). It follows that the n → π* transition (which is polarised perpendicular to the α-helix long axis) is at ~58° ($LD^r_{220\ nm} = 0.006$) from the average lipid direction. The success of this analysis is dependent on having a value for S (in this case from the retinal chromophore) and the protein being almost totally helical. With more varied protein structures, calculations to determine the net protein LD will be essential.

3.4 Protein LD calculations

It is apparent from the above discussion that LD data are consistent with proteins of known structure and a qualitative level of structural interpretation is possible. This can lead to mechanistic insights not possible from other techniques as in the case of FtsZ.[98] The level of interpretation that is possible, however, is restricted due to the cancellation of LD signals by overlapping bands as illustrated in Fig. 14. To get beyond this we need to know what each transition is doing and to be able to calculate spectra as a function of geometry until a good match between experiment and calculation is obtained.

If one is able to calculate the CD spectrum (see above), values for the components of the electric dipole transition moments are generated in the process. In the case of a single, perfectly oriented molecule we may write, using the notation of section 2.2

$$\mathrm{LD}_{\mathrm{molecule}} = (A_{||} - A_{\perp})_{\mathrm{molecule}} \propto \left\{ |\langle \psi^k | \mu_Z | \psi^0 \rangle|^2 - |\langle \psi^k | \mu_Y | \psi^0 \rangle|^2 \right\} \qquad (23)$$

where Z is the orientation axis and Y is perpendicular to Z. For the case of a perfectly oriented molecule Z is usually the longest axis of the molecule (denoted z), and Y is a rotational average over the molecular axes x and y. So the LD can be calculated using the same matrix methodology as used for CD, with an expectation of a more accurate match to experiment as LD depends only on the electric moments.

Unfortunately, samples are seldom if ever perfectly oriented and then we need to determine how to relate the components

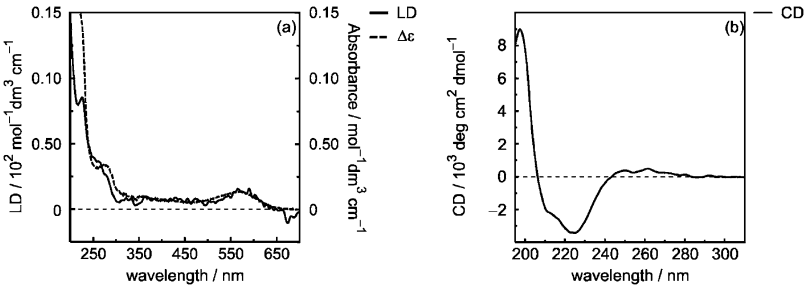


Fig. 16 Spectra of bacteriorhodopsin (0.2 mg/mL) added to a soya bean liposome solution (0.5 mg/mL). (a) Absorption (dashed line, 1 mm pathlength, baseline: liposome absorption spectrum) and LD (solid line, 0.5 mm pathlength, baseline: LD spectrum of sample without rotation spectrum); (b) CD (1 mm pathlength, baseline: liposome CD spectrum).

of the electric dipole transition moments calculated in the molecular axis system to those in the macroscopic system. Further, we need to consider an ensemble of molecules with their z axes distributed about the macroscopic orientation axis Z. In many cases, either the averaging inherent in the orientation method or the nature of the sample means that we can assume local uniaxial orientation (Fig. 17) in which case:

$$\mathrm{LD} = \frac{\mu^2}{2} S(3\cos^2\alpha - 1) \qquad (24)$$

or

$$\mathrm{LD}^r = \frac{\mathrm{LD}}{A} = \frac{A_\parallel - A_\perp}{A} = \frac{3}{2} S(3\cos^2\alpha - 1) \qquad (25)$$

where S may be viewed as a scaling factor defining the efficiency of macroscopic orientation. $A_\parallel$, $A_\perp$, and A all have $C\ell$ as a common factor. Hence, when using LD^r we do not need to know the sample concentration C or pathlength ℓ. Thus, $S = 1$ for perfect orientation (assumed in eqn (19)) and $S = 0$ for random (*i.e.* no) orientation. If the sample is macroscopically as well as molecularly uniaxial, such as molecules oriented in a polymer film drawn in one direction or polar molecules in an electric field, there is a simple relation between $A_\parallel$, $A_\perp$, and A which makes it unnecessary to measure all three quantities (Fig. 17).

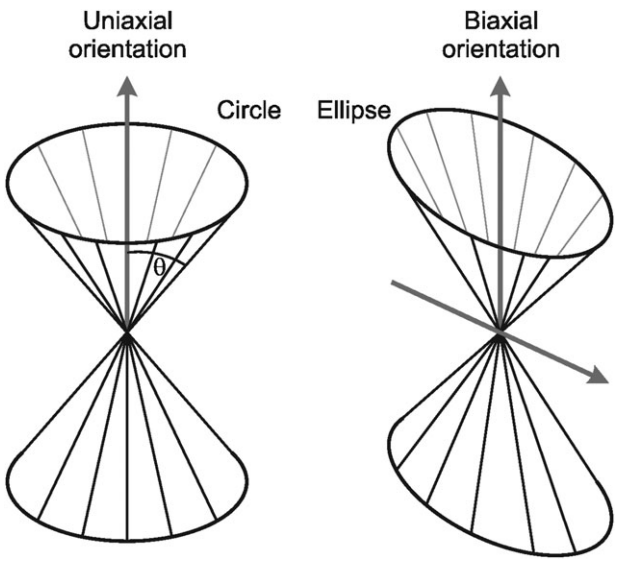


Fig. 17 Uniaxial and biaxial orientation. Uniaxial orientation requires all orientations on a cone about the orientation axis to be statistically equally probable. This may be due to static or dynamic factors.[90]

The challenge from a computational point of view is to integrate calculations of LD spectra for monomeric proteins of known structure with proposed polymerisation geometries and orientation distributions to enable LD experiment plus calculations to be used to determine fibre geometries. This may well involve molecular dynamics simulations to enable an estimate of S. To date LD calculations using the matrix method have been limited to calculating the expected LD for secondary structural motifs or chromophores.[90,102] The next stage is to tackle monomeric units of fibrous proteins, oriented as in a fibre or membrane proteins in liposomes, whose actual orientation cannot be easily determined in experiment. In general, the orientation dependence of an LD experiment is an additional parameter which needs to be taken into account in the calculation. Since even the best aligned liquid sample will never be perfectly oriented in reality (as this would be the case in a crystal), these deviations need to be considered in the calculation to get a more accurate simulation. In practice, this is done by calculating the perfectly aligned protein, including rotational averaging about its long axis. It is then tilted about a small angle and the LD spectra calculated. Comparison of spectra computed for different inclination angles with the experimental spectrum should allow the assignment of the most likely orientation of the protein. This methodology is currently being applied to several problems.

4. Online calculation using dichrocalc

The extensive use of the internet nowadays has prompted a large number of online services for protein analyses, *e.g.* for the interpretation of mass spectra, investigations into protein similarity and the prediction of secondary structure. Specifically concerning CD, the Protein Circular Dichroism DataBank (PCDDB) aims to build a freely available database of CD spectra[108] and the web interface DichroWeb provides analyses of CD data using several open source algorithms.[109] Dichroweb currently supports the programmes SELCON3,[110,111] CONTINILL,[112,113] CDSSTR,[114–116] VARSLC[114–116] and K2D[117] and combines seven different reference databases to choose from. Using these tools, it

calculates the secondary structure from the experimental spectrum, a goodness-of-fit parameter for the analyses and several overviews of calculation and experiments.

Although it is much more common for researchers to wish to extract structural information from a CD spectrum, it is sometimes of interest to generate a CD spectrum from a structure. If one suspects a protein contains a novel structure which the structure fitting databases do not include, then this approach is the only option. Alternatively, if the protein structure is dynamic and one wishes more information than a net average, then calculated spectra for different conformers are attractive. We have developed a freely accessible web interface implementing the matrix method. The required input is a PDB file to define the protein geometry. Files can also be retrieved from the RCSB Protein Data Bank[118] *via* their PDB code. An experimental spectrum for comparison may be included and a number of archive types can be uploaded to run calculations on several proteins.

After uploading a single PDB file or an archive containing many PDB files and/or giving PDB codes to fetch files from the RCSB, the user can choose whether to include the side chain contributions, charge transfer transitions, the curve type of the band shapes and the bandwidth used for them. Units of ellipticity or absorption may be used. If the NMAabinit parameter set is chosen, the user can also include three or four backbone transitions instead of the two selected by default. In addition to the CD spectrum, DichroCalc can also produce the LD spectrum, thus covering all the possibilities described in the preceding sections. After submitting the job, the results are returned by email and contain the coordinate files as well as the comparison plots of calculation and experimental data (if provided). Both the NMAabinit and NMAsemi parameter set are available for use in the calculation. The interface can be reached at http://comp.chem.nottingham.ac.uk/dichrocalc/. We hope that it will be of use to researchers undertaking studies similar to those described in section 2.6.

5. Conclusion

At the moment, the quality of CD spectra calculated by the matrix method depends very much on the structure of the protein in question. The CD spectra of highly α-helical proteins can be calculated with a Spearman rank correlation of 0.92 between the experimental and calculated intensity at 220 nm. Since the α-helix is the most common secondary structure type, the quantitative understanding of its contribution to the CD is a substantial achievement. The effects of random coil and particularly β-II structures on the observed CD spectra have to be better understood to achieve the level of agreement that can be reached for α-helical proteins. It is generally the case that, apart from the β-II class proteins, the *ab initio* set of parameters tends to overestimate CD intensities, especially in the region around 190 nm. The role of the various parameters in the calculations and the underlying assumptions of the matrix methodology are being further investigated. Furthermore, we are attempting to improve the details of the shape of the calculated spectra, *e.g.* to achieve the two minima of the α-helix between 208 and 220 nm without arbitrary choices of parameters to achieve the goal.

The significant progress which has been made in calculating CD spectra over the last few years is now being transferred over to LD spectroscopy. Protein LD data are becoming more available and it has become clear that they contain a great deal of structural information. However, due to overlapping transitions this information is often impossible to extract without understanding the spectroscopy of the system. The matrix method is directly applicable to LD data interpretation, as LD depends only on the electric moments of the transition, spectral shapes are easier to calculate. The issue of degree of sample orientation remains a challenge both for the experiment and the data interpretation.

Acknowledgements

We thank the Engineering and Physical Sciences Research Council (EPSRC) for funding (grant number GR/T09224) and Dr Daniel Barthel, University of Nottingham, for useful discussions.

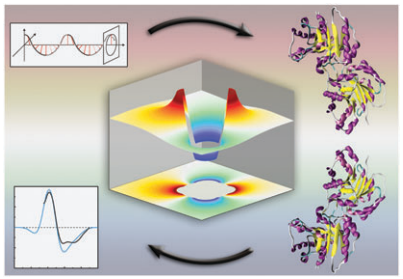
References

1 D.-F.-M. Arago, *Mem. Inst. Fr.*, 1811, **12**, 93.
2 J.-B. Biot, *Mem. Inst. Fr.*, 1812, **13**, 1–372.
3 J.-B. Biot, *Bull. Soc. Philomath.*, 1815, 190.
4 G. D. Fasman, *Circular Dichroism and the conformational analysis of biomolecules*, Plenum Press, New York, 1996.
5 L. Pasteur, *C. R. Acad. Sci. Paris*, 1848, **26**, 535–539.
6 J. A. Le Bel, *Bull. Soc. Chim. Fr.*, 1874, **22**, 337–347.
7 J. H. van't Hoff, *La Chimie dans L'Espace*, 1875.
8 G. Snatzke, *Angew. Chem., Int. Ed. Engl.*, 1968, **7**, 14–25.
9 N. Berova, K. Nakanishi and R. W. Woody, *Circular Dichroism: Principles and Applications*, John Wiley and Sons, New York, 2nd edn, 2000.
10 M. Born, *Phys. Z.*, 1915, **16**, 251–258.
11 W. Kuhn, *Z. Phys. Chem. B*, 1929, **4**, 14–36.
12 L. Rosenfeld, *Z. Phys.*, 1928, **52**, 161–174.
13 D. D. Fitts and J. G. Kirkwood, *Proc. Natl. Acad. Sci. U. S. A.*, 1956, **42**, 33–36.
14 W. Moffitt, *J. Chem. Phys.*, 1956, **25**, 467–478.
15 W. Moffitt, *Proc. Natl. Acad. Sci. U. S. A.*, 1956, **42**, 736–746.
16 A. S. Davydov, *Theory of molecular excitations*, Plenum Press, New York, 1971.
17 P. Urnes and P. Doty, *Adv. Protein Chem.*, 1961, **16**, 401–544.
18 G. Holzwarth and P. Doty, *J. Am. Chem. Soc.*, 1965, **87**, 218–228.
19 K. M. Specht, J. Nam, D. M. Ho, N. Berova, R. K. Kondru, D. N. Beratan, P. Wipf, R. A. Pascal and D. Kahne, *J. Am. Chem. Soc.*, 2001, **123**, 8961–8966.
20 K. Tanaka, Y. Itagaki, M. Satake, H. Naoki, T. Yasumoto, K. Nakanishi and N. Berova, *J. Am. Chem. Soc.*, 2005, **127**, 9561–9570.
21 P. Butz, G. E. Tranter and J. P. Simons, *PhysChemComm*, 2002, **5**, 91–93.
22 F. Furche, R. Ahlrichs, C. Wachsmann, E. Weber, A. Sobanski, F. Vogtle and S. Grimme, *J. Am. Chem. Soc.*, 2000, **122**, 1717–1724.
23 M. Schreiber, R. Vahrenhorst, V. Buss and M. P. Fulscher, *Chirality*, 2001, **13**, 571–576.
24 H. DeVoe, *J. Chem. Phys.*, 1964, **41**, 393–400.
25 H. DeVoe, *J. Chem. Phys.*, 1965, **43**, 3199–3208.
26 J. Applequist, *J. Chem. Phys.*, 1979, **71**, 4332–4338.
27 J. Applequist, K. R. Sundberg, M. L. Olson and L. C. Weiss, *J. Chem. Phys.*, 1979, **70**, 1240–1246.
28 B. K. Sathyanarayana and J. Applequist, *Int. J. Pept. Protein Res.*, 1985, **26**, 518–527.
29 K. A. Bode and J. Applequist, *J. Phys. Chem.*, 1996, **100**, 17825–17834.
30 K. L. Carlson, S. L. Lowe, M. R. Hoffmann and K. A. Thomasson, *J. Phys. Chem. A*, 2006, **110**, 1925–1933.

31 S. L. Lowe, R. R. Pandey, J. Czlapinski, G. Kie-Adams, M. R. Hoffmann, K. A. Thomasson and K. S. Pierce, *J. Pept. Res.*, 2002, **61**, 189–201.
32 P. M. Bayley, E. B. Nielsen and J. A. Schellman, *J. Phys. Chem.*, 1969, **73**, 228–243.
33 A. Rodger, J. Rajendra, R. Marrington, M. Ardhammar, B. Norden, J. D. Hirst, A. T. B. Gilbert, T. R. Dafforn, D. J. Halsall, C. A. Woolhead, C. Robinson, T. J. T. Pinheiro, J. Kazlauskaite, M. Seymour, N. Perez and M. J. Hannon, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4051–4057.
34 S. M. Kelly, T. J. Jess and N. C. Price, *Biochim. Biophys. Acta*, 2005, **1751**, 119–139.
35 L. A. Nafie, J. C. Cheng and P. J. Stephens, *J. Am. Chem. Soc.*, 1975, **97**, 3842–3843.
36 L. A. Nafie and M. Diem, *Acc. Chem. Res.*, 1979, **12**, 296–302.
37 T. A. Keiderling, *Appl. Spectrosc. Rev.*, 1981, **17**, 189–226.
38 L. D. Barron, *Molecular Light Scattering and Optical Activity*, Cambridge University Press, 1982.
39 G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner and A. Moscowitz, *J. Am. Chem. Soc.*, 1974, **96**, 251–252.
40 P. M. Bayley, *Prog. Biophys. Mol. Biol.*, 1973, **27**, 1–76.
41 M. T. Oakley, B. M. Bulheller and J. D. Hirst, *Chirality*, 2006, **18**, 340–347.
42 S. Brahms and J. Brahms, *J. Mol. Biol.*, 1980, **138**, 149–178.
43 C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry*, W. H. Freeman, San Francisco, 1980.
44 W. C. Johnson, *Proteins—Structure Function and Genetics*, 1990, **7**, 205–214.
45 P. Wallimann, R. J. Kennedy and D. S. Kemp, *Angew. Chem., Int. Ed. Engl.*, 1999, **38**, 1290–1292.
46 J. T. Yang, C. S. C. Wu and H. M. Martinez, *Meth. Enzymol.*, 1986, **130**, 208–269.
47 P. Z. Luo and R. L. Baldwin, *Biochemistry*, 1997, **36**, 8413–8421.
48 C. D. Andrew, S. Bhattacharjee, N. Kokkoni, J. D. Hirst, G. R. Jones and A. J. Doig, *J. Am. Chem. Soc.*, 2002, **124**, 12706–12714.
49 Z. Dang and J. D. Hirst, *Angew. Chem., Int. Ed.*, 2001, **40**, 3619–3621.
50 S. Bhattacharjee, G. Toth, S. Lovas and J. D. Hirst, *J. Phys. Chem. B*, 2003, **107**, 8682–8688.
51 K. W. Plaxco and C. M. Dobson, *Curr. Opin. Struct. Biol.*, 1996, **6**, 630–636.
52 J. D. Hirst, S. Bhattacharjee and A. V. Onufriev, *Faraday Discuss.*, 2003, **122**, 253–267.
53 E. L. Eliel, S. H. Wilen and L. N. Mander, *Stereochemistry of Organic Compounds*, Wiley, New York, 1994.
54 N. Sreerama and R. W. Woody, *Meth. Enzymol.*, 2004, **383**, 318–351.
55 I. Tinoco, *Adv. Chem. Phys.*, 1962, **4**, 113–160.
56 J. G. Kirkwood, *J. Chem. Phys.*, 1937, **5**, 479–491.
57 R. W. Woody, *J. Chem. Phys.*, 1968, **49**, 4797–4806.
58 R. W. Woody and I. Tinoco, *J. Chem. Phys.*, 1967, **46**, 4927–4945.
59 J. D. Hirst and N. A. Besley, *J. Chem. Phys.*, 1999, **111**, 2846–2847.
60 N. A. Besley and J. D. Hirst, *J. Am. Chem. Soc.*, 1999, **121**, 9636–9644.
61 R. W. Woody and N. Sreerama, *J. Chem. Phys.*, 1999, **111**, 2844–2845.
62 A. Bernhardsson, R. Lindh, G. Karlstrom and B. O. Roos, *Chem. Phys. Lett.*, 1996, **251**, 141–149.
63 G. Karlstrom, *J. Phys. Chem.*, 1989, **93**, 4952–4955.
64 G. Karlstrom, *J. Phys. Chem.*, 1988, **92**, 1315–1318.
65 L. Serrano-Andrés, M. P. Fülscher and G. Karlstrom, *Int. J. Quantum Chem.*, 1997, **65**, 167–181.
66 J. D. Hirst, K. Colella and A. T. B. Gilbert, *J. Phys. Chem. B*, 2003, **107**, 11813–11819.
67 Y. H. Chen, J. T. Yang and H. M. Martinez, *Biochemistry*, 1972, **11**, 4120–4131.
68 D. M. Rogers and J. D. Hirst, *J. Phys. Chem. A*, 2003, **107**, 11191–11200.
69 J. R. Platt, *J. Chem. Phys.*, 1949, **17**, 484–495.
70 Y. Yamamoto and J. Tanaka, *Bull. Chem. Soc. Jpn.*, 1972, **45**, 1362–1366.
71 D. M. Rogers and J. D. Hirst, *Biochemistry*, 2004, **43**, 11092–11102.
72 N. Sreerama and R. W. Woody, *Protein Sci.*, 2003, **12**, 384–388.
73 N. A. Besley and J. D. Hirst, *J. Mol. Struct. (THEOCHEM)*, 2000, **506**, 161–167.
74 W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Cambridge University Press, Cambridge, 1992.
75 R. A. Laskowski, M. W. Macarthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
76 J. C. Sutherland, E. J. Desmond and P. Z. Takacs, *Nucl. Instrum. Meth.*, 1980, **172**, 195–199.
77 R. W. Janes, *Bioinformatics*, 2005, **21**, 4230–4238.
78 B. A. Wallace and R. W. Janes, *Curr. Opin. Chem. Biol.*, 2001, **5**, 567–571.
79 K. Matsuo, R. Yonehara and K. Gekko, *J. Biochem.*, 2004, **135**, 405–411.
80 K. Matsuo, R. Yonehara and K. Gekko, *J. Biochem.*, 2005, **138**, 79–88.
81 A. J. Miles and B. A. Wallace, *Chem. Soc. Rev.*, 2006, **35**, 39–51.
82 M. T. Oakley and J. D. Hirst, *J. Am. Chem. Soc.*, 2006, **128**, 12414–12415.
83 M. B. Robin, *Higher Excited States of Polyatomic Molecules*, Academic Press, New York, 1975.
84 N. A. Besley and J. D. Hirst, *J. Phys. Chem. A*, 1998, **102**, 10791–10797.
85 A. P. Demchenko, *Ultraviolet Spectroscopy of Proteins*, Springer, Berlin, 1986.
86 R. Weinkauf, P. Schanen, D. Yang, S. Sonkara and E. W. Schlag, *J. Phys. Chem.*, 1995, **99**, 11255–11265.
87 W. G. McGimpsey, L. Chen, R. Carraway and W. N. Samaniego, *J. Phys. Chem. A*, 1999, **103**, 6082–6090.
88 L. Serrano-Andrés and M. P. Fülscher, *J. Phys. Chem. B*, 2001, **105**, 9323–9330.
89 A. T. B. Gilbert and J. D. Hirst, *J. Mol. Struct. (THEOCHEM)*, 2004, **675**, 53–60.
90 A. Rodger and B. Nordén, *Circular Dichroism & Linear Dichroism*, Oxford University Press, Oxford, 1997.
91 M. Couette, *Ann. Chim. Phys.*, 1890, **6**, 433–510.
92 A. Mallock, *Proc. R. Soc. London*, 1888, **45**, 126.
93 A. Rodger, R. Marrington, M. A. Geeves, M. Hicks, L. de Alwis, D. J. Halsall and T. R. Dafforn, *Proc. R. Soc. London*, 2006, **8**, 3161–3171.
94 S. Higashi, M. Kasai, F. Oosawa and A. Wada, *J. Mol. Biol.*, 1963, **7**, 421–430.
95 A. Wada, *Appl. Spectrosc. Rev.*, 1972, **6**, 1–30.
96 T. R. Dafforn, J. Rajendra, D. J. Halsall, L. C. Serpell and A. Rodger, *Biophys. J.*, 2004, **86**, 404–410.
97 T. R. Dafforn and A. Rodger, *Curr. Opin. Struct. Biol.*, 2004, **14**, 541–546.
98 R. Marrington, E. Small, A. Rodger, T. R. Dafforn and S. Addinall, *J. Biol. Chem.*, 2004, **47**, 48821–48829.
99 R. Marrington, M. Seymour and A. Rodger, *Chirality*, 2006, in press.
100 R. Marrington, T. R. Dafforn, D. J. Halsall, J. I. MacDonald, M. Hicks and A. Rodger, *Analyst*, 2005, **130**, 1608–1616.
101 M. Ardhammar, N. Mikati and B. Nordén, *J. Am. Chem. Soc.*, 1998, **120**, 9957–9958.
102 A. Rodger, J. Rajendra, R. Marrington, R. Mortimer, T. Andrews, J. B. Hirst, A. T. B. Gilbert, D. Halsall, T. Dafforn, M. Ardhammar, B. Nordén, C. A. Woolhead, C. Robinson, T. Pinheiro, K. J. M. Seymour, N. Perez and M. J. Hannon, in *Biophysical Chemistry: Membranes and Proteins*, ed. R. H. Templer and R. Leatherbarrow, The Royal Society of Chemistry, Cambridge, Editon edn., 2002, pp. 3–19.
103 J. Rajendra, A. Damianoglou, M. Hicks, P. Booth, P. M. Rodger and A. Rodger, *Chem. Phys.*, 2005.
104 J. P. Cartailler and H. Luecke, *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **32**, 285–310.
105 W. Meijberg and P. J. Booth, *J. Mol. Biol.*, 2002, **319**, 839–853.
106 D. Oesterhelt and L. Schuhmann, *FEBS Lett.*, 1974, **44**, 262–265.
107 B. Albinsson, M. Kubista, B. Nordén and E. W. Thulstrup, *J. Phys. Chem.*, 1989, **93**, 6646–6654.
108 L. Whitmore, B. A. Wallace and R. W. Janes, *The Protein Circular Dichroism DataBank (PCDDB)*; http://pcddb.cryst.bbk.ac.uk/, 2006.
109 L. Whitmore and B. A. Wallace, *Nucleic Acids Res.*, 2004, **32**, W668–W673.

110 N. Sreerama, S. Y. Venyaminov and R. W. Woody, *Protein Sci.*, 1999, **8**, 370–380.
111 N. Sreerama and R. W. Woody, *Anal. Biochem.*, 1993, **209**, 32–44.
112 S. W. Provencher and J. Glockner, *Biochemistry*, 1981, **20**, 33–37.
113 I. H. M. Vanstokkum, H. J. W. Spoelder, M. Bloemendal, R. Vangrondelle and F. C. A. Groen, *Anal. Biochem.*, 1990, **191**, 110–118.
114 L. A. Compton and W. C. Johnson, *Anal. Biochem.*, 1986, **155**, 155–167.
115 P. Manavalan and W. C. Johnson, *Anal. Biochem.*, 1987, **167**, 76–85.
116 N. Sreerama and R. W. Woody, *Anal. Biochem.*, 2000, **287**, 252–260.
117 M. A. Andrade, P. Chacon, J. J. Merelo and F. Moran, *Protein Eng.*, 1993, **6**, 383–390.
118 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.

Physical Chemistry Chemical Physics b615870f



Q1 **Circular and linear dichroism of proteins†**

Benjamin M. Bulheller, Alison Rodger and Jonathan D. Hirst

We review methods for calculating protein CD spectra and present a new web interface for online CD and LD calculations.

Please check this proof carefully. **Our staff will not read it in detail after you have returned it.** Translation errors between word-processor files and typesetting systems can occur so the whole proof needs to be read. Please pay particular attention to: tabulated material; equations; numerical data; figures and graphics; and references. If you have not already indicated the corresponding author(s) please mark their name(s) with an asterisk. Please e-mail a list of corrections or the PDF with electronic notes attached -- do not change the text within the PDF file or send a revised manuscript.

Please bear in mind that minor layout improvements, e.g. in line breaking, table widths and graphic placement, are routinely applied to the final version.

Please note that, in the typefaces we use, an italic vee looks like this: ν, and a Greek nu looks like this: ν.

We will publish articles on the web as soon as possible after receiving your corrections; **no late corrections will be made.**

Please return your **final** corrections, where possible within **48 hours** of receipt, by e-mail to: proofs@rsc.org

Reprints—Electronic (PDF) reprints will be provided free of charge to the corresponding author. Enquiries about purchasing paper reprints should be addressed via: http://www.rsc.org/Publishing/ReSourCe/PaperReprints/. Costs for reprints are below:

Reprint costs		
No of pages	Cost for 50 copies	Cost for each additional 50 copies
2-4	£180	£115
5-8	£300	£230
9-20	£600	£480
21-40	£1100	£870
>40	£1700	£1455
Cost for including cover of journal issue: £50 per 50 copies		

Queries are marked on your proof like this Q1, Q2, etc. and for your convenience line numbers are indicated like this 5, 10, 15, ...

Query reference	**Query**	**Remarks**
Q1	For your information: You can cite this article before you receive notification of the page numbers by using the following format: (authors), (journal), (year), DOI: 10.1039/ (paper number).	

Influence of surface shape on DNA binding of bimetallo helicates

Jemma C. Peberdy,[a] Jaroslav Malina,[b] Syma Khalid,[c] Michael J. Hannon,[†d] Alison Rodger[†a]

(a) Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.
(b) Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic
(c) Department of Biochemistry, University of Oxford, Oxford, OX1 1QU, UK
(d) School of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

‡

Abstract

In order to probe the DNA-helicate interactions responsible for the DNA binding and remarkable changes of the DNA secondary structure induced by a tetracationic bi-metallo helicate $[Fe_2(L^1)_3]^{4+}$ ($L^1=C_{25}H_{20}N_4$), we have designed and synthesised derivatives with hydrophobic methyl groups at different positions on the ligand backbone. Two dimetallo helicates $[Fe_2(L^i)_3]^{4+}$ were prepared using ligands L^3 and L^5 with the methyl substituent on, respectively the 3 and 5 positions of the pyridyl ring thus producing a wider or slightly longer tetracationic DNA binder. UV/visible absorbance, circular and linear dichroism spectroscopies have been used to characterize the interactions of the cylinders with DNA with the aim of investigating any sequence preference or selectivity upon binding. Competitive binding studies using fluorescent dyes Hoechst 33258 (a minor groove binder), ethidium bromide (an intercalator) and a major groove binding cation (cobalt (III) hexammine) which induces the B→Z transition have been employed to determine the binding geometries of the enantiomers of two methylated helicates (L^3 and L^5) to DNA and compare with the data obtained previously for the unmethylated analogue (L^1). The results demonstrate that the racemic mixtures and the resolved enantiomers of all helicates bind to DNA inducing structural changes. The overall conclusion from the effect of adding these groups to the surface of the parent helicate is that increasing the width (L^3) reduces the DNA binding strength, the bending and coiling effect and the groove selectivity of the enantiomers compared with the parent compound. There is limited evidence to suggest a slight GC sequence preference. Lengthening the helicate (L^5) results in DNA interactions similar to those of the parent compounds, with an increased preference of the P-enantiomer for the minor groove indicating an enhancement of mode selectivity.

† Corresponding authors: M.J. Hannon: Phone: 44-121-414 2527, email: m.j.hannon@bham.ac.uk;A. Rodger: Phone: 44-24-76574696, email: a.rodger@warwick.ac.uk, fax: 44-24-76575795.

‡ Dedicated to the memory of Edward I. Stiefel

Introduction

Agents that can bind to DNA and regulate its expression find important application both for as probes in fundamental biological studies and in the clinic for disease control and treatment.[1-4] In particular, agents that bind to a broad spectrum of sites on DNA are in very wide-spread clinical use for treating cancers and viral diseases in both humans and other mammals. Such drugs include *minor*-groove binding agents (such as berenil),[5] intercalators (such as anthracyclines) [6] and agents that form direct bonds to the nitrogens of the DNA bases (such as the cisplatin family of anti-cancer drugs).[7-9] While these drugs are extremely successful there are many problems still to be addressed such as limitations in their activity against some cancers, and acquired resistance; a key feature for creating future generations of drugs is to investigate and understand the action of agents that bind to DNA through different molecular-level actions, since this will be crucial to gain different clinical activities and to circumvent potential cross-resistance with current drugs.

We have recently described a new class of synthetic metallo-supramolecular helicates which bind strongly and non-covalently to the *major*-groove of DNA and cause remarkable and unprecedented effects on DNA structure, coiling-up the DNA in a manner somewhat reminiscent of histones.[10-13] The helicates are tetracations and are comprised of three organic ligand strands wrapped in a helical fashion about two iron(II) metal centres, with face-edge π-π as well as the metal-ligand interactions contributing to the structural rigidity of the unit (Fig 1).[14] Despite their cationic nature, their surfaces are predominantly comprised of hydrocarbon units from aryl rings. As illustrated in Fig 1, the complexes are chiral *tris*-chelate systems. By convention the left handed or negatively twisted helicate is labelled (M) and the right handed or positively twisted helicate labelled (P). The two enantiomers of the helicates (P and M) can be resolved on cellulose [15] (which itself possesses chiral grooves on its surface) and the M enantiomer has been shown to be more aggressive than the P enantiomer in the extent of induced DNA coiling. We have speculated that this might be a consequence of subtle differences in their binding.[11, 16] We have shown that introducing methyl substituents onto the surface of the helicate can affect the relative bindings of the P and M enantiomers to cellulose.[17] We reasoned that it might similarly affect the binding to DNA and that such changes might provide information on the subtle differences in the DNA binding modes of the unsubstituted enantiomers which might then underpin a programme both to control the DNA binding and to explore routes to achieve sequence selective binding. Herein we describe the DNA binding of the enantiomers of two different hexa-methyl substituted helicates (Fig 1).

Through out this paper we use R-L^3 for example to refer to the racemic mixture of the helicate made from the ligand with methyl substituents on the 3-position. UV/visible absorbance, circular and linear dichroism have been used to investigate the interactions of the metal complex with DNA with the aim of investigating any sequence preference or selectivity upon binding. Competitive binding studies using fluorescent dyes (Hoechst 33258 and ethidium bromide) and a major groove binding cation (cobalt (III) hexammine) which induces the B→Z transition have been used to probe the binding geometries of the enantiomers of two methylated helicates to DNA.

Material and methods

The dimetallo helicates $[Fe_2(L^i)_3]^{4+}$ made using ligands L^1 (no methyl subsituents), L^3 and L^5 (with the methyl substituent on, respectively the 3 and 5 positions of the pyridyl ring, Fig 1) were synthesised and resolved according to methods described in refs.[14, 15, 17] Stock aqueous solutions of the racemic and enantiomeric helicates were prepared on the day of the experiment from the solid chloride salt and the concentrations determined using UV/visible absorbance spectroscopy and the extinction coefficient of the parent compound ($\varepsilon_{574\,nm}$ = 16,900 M^{-1} cm^{-1}) for the longest wavelength maximum of each compound (see Fig 2) as described in [ref. 17]. All DNA-binding experiments were carried out in sodium cacodylate buffer (pH 6.8, 1 mM) and NaCl (50 mM) which was sufficient to keep the DNA and helicates in solution throughout titration series. UV/visible absorbance and *CD* (circular dichroism) and *LD* (linear dichroism) spectra were recorded on Jasco V-550 and J-715 spectrophotometers respectively using a 1 cm path length cuvette. Black quartz 4 mm wide cuvettes were used to reduce the volume of sample required. A water baseline was subtracted from each data set. The samples have no significant absorbance intensity at 750 nm, however, the *CD* spectrum is not zero at this wavelength (presumably an electric dipole forbidden, magnetic dipole allowed *d-d* transition which has small but non-zero *CD* intensity). Thus the *CD* spectra were zeroed at 420 nm for all experiments. Fluorescence experiments were carried out using a Perkin Elmer LS50B spectrometer. All experiments were performed in duplicate to ensure reproducibility.

Calf thymus DNA (ct-DNA) was purchased from Sigma Aldrich. Synthetic double stranded DNA alternating homopolymers, poly[d(A-T)$_2$] (AT) and poly[d(G-C)$_2$] (GC) were obtained from Pharmacia Biochemicals. Ultrapure water (18.2 ΩM) was used in all experiments. All polynucleotides were dissolved in water without any further purification and kept frozen until the day of the experiment. Polynucleotide concentrations were determined spectrophotometrically using the molar extinction coefficients:[18, 19] ct-DNA ε_{260} = 6,600 M^{-1} cm^{-1}, AT-DNA ε_{262} = 6,600 M^{-1} cm^{-1} and GC-DNA ε_{254} = 8,400 M^{-1} cm^{-1}. A 100 mM stock solution of sodium cacodylate buffer was prepared by mixing 50 mL of 0.2 M sodium cacodylate (4.28 g of $Na(CH_3)_2AsO_2.3H_2O$ in 100 mL) with 9.3 mL of 0.2 M hydrochloric acid, and diluting to a total volume of 200 mL. Ethidium bromide, Hoechst 33258 and $[Co(NH_3)_6]^{3+}$ were purchased from Sigma-Aldrich, dissolved in water and used without any further purification. The ethidium bromide concentration was determined spectrophotometrically using the molar extinction coefficient ε_{480} = 5,600 M^{-1} cm^{-1},[20] and those of Hoechst 33258 and $[Co(NH_3)_6]^{3+}$ by weight.

DNA binding studies: Absorbance and *CD* spectral data as a function of concentration were collected by preparing a starting solution of DNA and helicate in buffer usually at ratio 30 base:1 helicate and then diluting the DNA concentration by adding a solution of helicate (at the original concentration) in buffer. Solutions of each enantiomer were stored in the fridge at 4 °C to prevent racemisation and the laboratory air conditioning was set to 18 °C. *LD* experiments by way of contrast were undertaken at constant DNA concentration (500 μM) and viscosity by adding aliquots of 1000 μM DNA when helicate was added. Helicate concentrations were kept below 50 μM to avoid precipitation of the DNA.

Competition assays: The competition assays were all undertaken with fixed DNA and competitor (ethidium bromide, Hoechst, $[Co(NH_3)_6]^{3+}$) concentrations and variable helicate. In general the experiments were designed so that the weaker binder was displaced by the stronger one.

Results

The bimetallo helicate absorption spectra show bands at ~580 nm (characteristic of the iron (II) metal to ligand charge transfer (MLCT) transitions); 320 nm (due to the imine bond in-ligand transitions); and in the DNA region of the spectrum (other in-ligand transitions) as illustrated in Fig 2. Thus the changes in UV/visible absorbance spectra and *CD* spectra (*ICD* are shown) for the helicates (see Figs 3) in different regions of the spectrum relate to different transitions.

It is apparent from the absorbance, *CD* and *LD* spectra of Figs 3 and 4 that R-, M-, and P-L^3 and L^5 all interact with ct-DNA. A bisignate signal is observed in the MLCT region of the induced *CD* (*ICD*) spectra of (R) and (M) helicates that is positive at longer wavelength and negative at shorter wavelength. (P)-L^3 shows a similar shape but less intensity and (P)-L^5 has almost no *ICD* in the MLCT region, though some in the in-ligand region. All helicates bend the DNA as evidenced by a decrease in the magnitude of the DNA *LD* band [10, 11] (Table 1) and have a positive MLCT *LD* signal. The in-ligand region *LD* is more variable and typically of small intensity — consistent with the mixed polarisation of these transition [21] so that small angle variations in binding geometry can lead to sign changes. It can be observed from these spectra that the extent of orientation of the metal complex is less significant for the enantiomers of the L^5 helicate than for the racemic mixture, as observed by a signal of lower intensity in the MLCT region of the spectrum. The overlay of the *CD* spectra (at 10:1 base pairs to helicate) for the enantiomers for each complex and the average of the (M) and (P) spectra (Fig 5) suggests that in the racemic mixture more (P)-L^3 is binding than (M)-L^3. With the L^5 helicates, the racemic mixture is (not surprisingly given the absence of a (P)-L^5 signal) dominated by that of the (M) enantiomer.

Whether the L^3 and L^5 helicates have any sequence preference was probed by measuring their induced *CD* spectra in the presence of AT and GC DNAs. Little variation in spectral shape over a mixing ratio range was observed. The data at 20:1 base:helicate mixing ratio are given in Fig 6. The spectrum obtained for (M)-L^3 in the presence of AT- differs from ct-DNA and GC suggesting when the complex has both AT and GC sequences available, as in ct-DNA, it prefers GC sequences. It is quite possible that a different binding mode is adopted with AT rich DNA. (P)-L^3 *ICD* with ct-DNA appears to be (very small signals) different from the two alternating homopolymers, suggesting different interactions with alternating and non-alternating DNAs. By way of contrast the spectra obtained for the L^5 helicates seem to show no sequence dependence as was the case for the L_1 helicates.[11]

Competition binding assays

Although simple electrostatics will ensure that all DNA binders compete with each other, if two molecules are targeting the same position on DNA the competition may be more marked. Thus, a series of competitive binding experiments using ethidium bromide (EB, an intercalator), Hoechst 33258 (a minor groove binder spanning 4–5 base pairs with a preference for A-T rich sequences) [18, 20, 22] and $[Co(NH_3)_6]^{3+}$ (a major groove binding cation which converts G-rich B-DNA to Z-DNA) [23] were designed to investigate the binding properties of the enantiomers of the metal complexes to DNA. EB and Hoechst (Fig 7) have fluorescence signals which are significantly enhanced upon DNA binding. It should be noted, however, that in the absence of helicates the probes are either bound or free, whereas in the presence of the helicates they may also interact with helicates on or off the DNA. In addition binding modes can changes.

Ethidium bromide displacement: The decrease in fluorescence intensity upon displacement of the EB from ct-DNA by the L^1 (the parent helicate without methyl substituents), L^3 and L^5 helicates at the wavelength maximum (590 nm) is shown in Fig 8. All of the helicates displace EB, so their binding constants are all higher than that of EB. The qualitative ranking of the helicate's ability to displace EB is given in Table 1. Quantification of the binding constants (following Appendix 1) was difficult because free EB has a higher fluorescence intensity at 590 nm than the fluorescence intensity observed at the end point of the titration series (when we assume all of it has been displaced). Implicit in this is an interaction between helicate with EB or helicate-DNA with EB which causes quenching of the EB emission when it is displaced from its intercalated binding site. The binding constant of each helicate, K_h, and the number of DNA bases it spans, n_h, determined using the equations of Appendix 1 given in Table 1 are therefore only indicative. Since K_h and n_h are not truly independent the binding constants for different helicates cannot be directly compared unless the site sizes are similar.

Hoechst 33258 competition binding assay: The fluorescence data of Fig 9 show the displacement of Hoechst 33258 upon addition of the helicates. The first observation is that all the helicates displace Hoechst to a similar extent. This was not expected as, for example, (M)-L^1 has been shown to bind in the major groove of B-DNA [10, 11, 24] while the binding mode of (P)-L^1 is less clear. The result may relate to the conformational change (as shown by the *LD* data in Figs 3 and 4) the helicates induce removing Hoechst's binding site. The displacement ranking is summarised in Table 1.

Cobalt (III) hexammine competitive binding assay: Circular dichroism was used to probe the transition from B→Z-DNA upon addition of $[Co(NH_3)_6]^{3+}$ to a GC or [helicate + GC] solution. The rationale behind these experiments is that B-DNA will be converted to Z-DNA only if the $[Co(NH_3)_6]^{3+}$ binds in the major groove. Fig 10 shows the B→Z transition of DNA (290 nm region goes from positive to negative) upon addition of increasing concentrations of $[Co(NH_3)_6]^{3+}$. The spectra of Figs 10 b and c show the least and most effective helicate for preventing the conversion to Z-DNA. Fig 11 is a summary of the data for all helicates. The rankings are summarised in Table 1.

Discussion and conclusions

The DNA binding modes of the enantiomers of two methylated ligands (L^3 and L^5) which form triple helicates have been studied by a variety of spectroscopic techniques and compared with the date obtained herein and previously [10, 11] for the unmethylated ligand L^1 (Fig 1). The racemic mixtures and the resolved enantiomers of all helicates bind to DNA inducing structural changes. The racemic mixtures have properties intermediate between those of the two enantiomers. The spectroscopy of DNA of the L^3 and (M)-L^5 helicates are similar to that of L^1. There is evidence of an element of GC selectivity with the (M)-L^3 enantiomer. In general, the binding strength as assessed by displacement of other DNA binders, of the L^3 helicates is the least of the three sets, which may account for its lower bending influence. If anything (P)-L^3 is more effective at displacing the major groove binder and (M)-L^3 the minor groove binder — in contrast to our expectations from the behaviour of the L^1 helicates. In fact (P)-L^3 did not displace all of the Hoechst minor groove binder from the DNA even at high loading.

L^5 helicates, like the parent, show no significant sequence preference (GC versus AT) and similar binding constants to those of L^1 and each other. This is not entirely surprising when one realises that the substituent methyl groups mainly serve to make a small extension to the length of the helicate and do not dramatically change the surface that interacts with the DNA bases. (P)-L^5 does, however, have very little MLCT induced *CD* compared with the other helicates, which suggests it interacts less with the DNA bases. Intriguingly (P)-L^5 is the P helicate which causes the most DNA bending (almost as effective as (M)-L^1 and L^5) and it displaces ethidium bromide and Hoechst fairly effectively while having less impact on the major groove binding $[Co(NH_3)_6]^{3+}$. We previously proposed competing bending induced by major and minor groove binding.[10, 11] So this is consistent with the increased minor groove selectivity of (P)-L^5 relative to (P)-L^3 and (P)-L^1. Conversely (M)-L^5 displaces $[Co(NH_3)_6]^{3+}$ the most effectively, has almost least effect on Hoechst, and is an effective DNA-bender. This would tend to suggest a major groove preference for (M)-L^5 and a minor groove, but not deep (thus no sequence preference), preference for (P)-L^5.

The displacement results for L^1 are also of interest, adding to our previous spectroscopic investigations of this compound. (M)-L^1 is very effective at displacing $[Co(NH_3)_6]^{3+}$ from the major groove and the least effective at displacing the minor groove binder Hoechst (similar behaviour to (M)-L^5). This is consistent with the major groove binding mode we have described for this compound.[10, 11, 24] The location of (P)-L^1 was more ambiguous and the displacement data reflects this. It is the most effective helicate at displacing the minor groove binder Hoechst but it is also effective as displacing $[Co(NH_3)_6]^{3+}$ from the major groove. It seems likely that this helicate may bind both in the major and minor grooves resulting in less effective net bending.

The overall conclusion from the effect of adding what is essentially just size to the surface of the parent helicate is that increasing the width (L^3) weakens the DNA binding strength, reduces the bending and coiling effect and also reduces the groove selectivity of the enantiomers compared with the parent compound. There is limited evidence to suggest a slight GC sequence preference. Lengthening the helicate by adding methyl groups (L^5) results in DNA interactions similar to those of the parent compounds, though there is an increased relative preference of the P-enantiomer for the minor groove indicating an enhancement of mode selectivity. (P)-L^5 is the only P helicate of the three which is an effective DNA bender. This may simply be that its extra length facilitaes bending, or it may be that the minor-groove induced bending opposes the major-groove bending so reducing the major groove occupancy leads to significant net bending.

Acknowledgements We thank the EU (MARCY RTN HPRN-CT-2002-00175; Marie Curie fellowship HPMF-CT-2002–01856) for support.

Appendix 1. Supplementary material

Calculation of a helicate DNA-binding constant from ethidium bromide displacement fluorescence data is provided as Supplementary material. Supplementary data associated with this article can be found, in the online version, at doi: (blank).

References

1. Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S., *Topics Current Chem.* **2005,** 253, 1-31.

2. Da Ros, T.; Spalluto, G.; Prato, M.; Saison-Behmoaras, T.; Boutorine, A.; Cacciari, B., Oligonucleotides and oligonucleotide conjugates: a new approach for cancer treatment. *Curr. Med. Chem.* **2005,** 12, (1), 71-88.
3. Hannon, M. J., Supramolecular DNA recognition. *Chem Soc Rev* **2007,** 36, (2), 280-95.
4. Richards, A. D.; Rodger, A., Synthetic metallomolecules as agents for the control of DNA structure. *Chemical Society Reviews* **2007,** 36, 471-483.
5. Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R., DNA minor groove binders as potential antitumor and antimicrobial agents. *Med. Res. Rev.* **2004,** 24, (4), 475-528.
6. Martinez, R.; Chacon-Garcia, L., The search of DNA-intercalators as antitumoral drugs: what it worked and what did not work. *Curr Med Chem* **2005,** 12, (2), 127-51.
7. Lippert, E. B., *Cisplatin, Chemistry and Biochemistry of a Leading Anti-Cancer Drug*. Wiley-VCH: Zürich, 1999.
8. Reedijk, J., Improved understanding in platinum antitumour chemistry. *Chemical Communications* **1996**, (7), 801-806.
9. Guo, Z. J.; Sadler, P. J., Medicinal inorganic chemistry. *Advances in Inorganic Chemistry* **2000,** 40, 183-306.
10. Hannon, M. J.; Moreno, V.; Prieto, M. J.; Moldrheim, E.; Sletten, E.; Meistermann, I.; Isaac, C. J.; Sanders, K. J.; Rodger, A., Intramolecular DNA Coiling Mediated by a Metallo-Supramolecular Cylinder Support by the Leverhulme Trust (F/215/BC) and the EPSRC lifesciences interface network (GR/M91105) is gratefully acknowledged. Discussions with Julie MacPherson have been of great assistance during preparation of the manuscript. *Angew Chem Int Ed Engl* **2001,** 40, (5), 879-884.
11. Meistermann, I.; Moreno, V.; Prieto, M. J.; Molderheim, E.; Sletten, E.; Khalid, S.; Rodger, P. M.; Peberdy, J.; Isaac, C. J.; Rodger, A.; Hannon, M. J., Intramolecular DNA coiling mediated by metallo-supramolecular cylinders: differential binding of P and M helical enantiomers. *Proceedings of the National Academy of Sciences* **2002,** 99, 5069-5074.
12. Uerpmann, C.; Malina, J.; Pascu, M.; Clarkson, G. J.; Moreno, V.; Rodger, A.; Grandas, A.; Hannon, M. J., "Design and DNA binding of an extended triple-stranded metallo-supramolecular cylinder". *Chemistry - A European Journal* **2005**.
13. Childs, L. J.; Malina, J.; Pascu, M.; Rolfsnes, B. E.; Prieto, M. J.; Broome, M. J.; Rodger, P. M.; Sletten, E.; Moreno, V.; Rodger, A.; Hannon, M. J., A DNA-binding cuprous metallo-supramolecular cylinder that can act as an artificial nuclease. *Chemistry: A European Journal* **2006,** 12, 4919–4927.
14. Hannon, M. J.; Painting, C. L.; Jackson, A.; Hamblin, J.; Errington, W., An inexpensive approach to supramolecular architecture. *Chemical Communications* **1997**, (18), 1807-1808.
15. Hannon, M. J.; Meistermann, I.; Isaac, C. J.; Blomme, C.; Aldrich-Wright, J. R.; Rodger, A., Paper: a cheap yet effective chiral stationary phase for chromatographic resolution of metallo-supramolecular helicates. *Chemical Communications* **2001**, (12), 1078-1079.
16. Peberdy, J. C.; Malina, J.; Khalid, S.; Hannon, M. J.; Rodger, A., Influence of surface shape on DNA binding of bimetallo helicates. *Journal of Inorganic Biochemistry* **2007**.
17. Kerckhoffs, J. M. C. A.; Peberdy, J. C.; Meistermann, I.; Childs, L. J.; Isaac, C. J.; Pearmund, C. R.; Reudegger, V.; Alcock, N. W.; Hannon, M. J.; Rodger, A., Enantiomeric resolution of supramolecular helicates with different surface topographies. *Dalton Transactions* **2007**, 734-742.
18. Waring, M. J., *J. Mol. Biol.* **1965,** 13, 269 – 282
19. Wells, R. D.; Larson, J. E.; Grant, R. C.; Shortle, B. E.; Cantor, C. R., Physicochemical studies on polydeoxyribonucleotides containing defined repeating nucleotide sequences. *J. Mol. Biol.* **1970,** 54, 465-497.
20. Mahler, H. R.; Kline, B.; Mehrota, B. D., *J. Mol. Biol.* **1964,** 9, 801-811.
21. Rodger, A.; Sanders, K. J.; Hannon, M. J.; Meistermann, I.; Parkinson, A.; Vidler, D. S.; Haworth, I. S., DNA structure control by polycationic species: polyamine, cobalt ammines, and di-metallo transition metal chelates. *Chirality* **2000,** 12, (4), 221-36.
22. Pjura, P. E.; Grzeskowlak, K.; Dickerson, D. E., *J. Mol. Biol.* **1987,** 197, 257-271
23. Parkinson, A.; Hawken, M.; Hall, M.; Sanders, K. J.; Rodger, A., Amine induced Z-DNA in poly(dG-dC).(dG-dC): circular dichroism and gel electrophoresis study. *PCCP* **2000,** 2, 5469-5478.
24. Moldrheim, E.; Hannon, M. J.; Meistermann, I.; Rodger, A.; Sletten, E., Interaction between a DNA oligonucleotide and a dinuclear iron(II) supramolecular cylinder; an NMR and molecular dynamics study. *J Biol Inorg Chem* **2002,** 7, (7-8), 770-80.
25. Marrington, R.; Dafforn, T. R.; Halsall, D. J.; Rodger, A., Micro-volume Couette flow sample orientation for absorbance and fluorescence linear dichroism. *Biophysical Journal* **2004,** 87, (3), 2002-2012.

Table 1 Order of effectiveness of displacement and bending data (as represented by % loss of 260 nm *LD* signal at 50:1 DNA base to helicate) summary for the helicates. ~ denotes equivalent to. Also shown are the approximate binding constant K_h and the metal complex site size n_h calculated using the methodology of Appendix 1. Errors in K_h (as determined by experiment to experiment variation) are of the order 20 – 50%. P1 denotes P-(L^1) *etc*.

	Least								**Most**
EB displacement	P3	R3 ~	M3	M1~	R1~	***M5***	**P5**~	R5	P1
K_h/ (10^6 M^{-1})	2	22	30	70	60	100	35	80	70
n_h / bases	20	10	20	25	130	30	45	25	55
% Loss in *LD*	P1	P3	M3	R5	R3	R1	**P5**	***M5***	M1
	20	25	35	40	40	60	65	70	70
Hoechst 33258 displacement	M1	***M5***	P3	R5	R3	M3	**P5**	R1	P1
$[Co(NH_3)_6]^{3+}$ displacement	R3	R5	**P5**	M3	P3	P1	R1	M1	***M5***
Sequence dependence									M3

L^1 helicate *LD* from Reference [11].

Fig 1 The chemical structure of the ligands (top), liqourice representation of the crystal structure of the M-parent helicate (middle), and the resulting shapes of M-L^3 (bottom left) and M-L^5 (bottom right) helicates.

Fig 2 UV/visible absorbance spectra of ct-DNA (200 μM), (R)-L^3 (17 μM) and (R)-L^5 (25 μM). Spectra collected in a 1 cm path length cuvette.

Fig 3 UV/visible absorbance (A, D, and G), *ICD* spectra (B, E, and H) and *LD* spectra (C, F and I) of (a) (R)-L^3, (b) (M)-L^3 and (c) (P)-L^3 binding to ct-DNA. Helicate identity is indicated on the absorbance figure for each row. The DNA base:helicate ratio for absorbance and ICD are indicated on the absorbance figures, that for the *LD* are shown on the *LD* figures. Absorbance and *CD* spectra were at 20 μM helicate with DNA concentrations to ensure the ratios indicated on the figure in the absorbance insert. A 1 cm path length cuvette was used, a water baseline was subtracted from each sample spectrum and the data zeroed at 420 nm. *LD* spectra were recorded at 500 μM DNA concentration and helicate concentrations to ensure the ratios indicated in the Figure. A 1 mm pathlength Couette flow cell[25] was used, a water baseline was subtracted from each sample spectrum and the data zeroed at 420 nm.

Fig 4 UV/visible absorbance (A, D, and G), *ICD* spectra (B, E, and H) and *LD* spectra (C, F and I) of (a) (R)-L^5, (b) (M)-L^5 and (c) (P)-L^5 binding to ct-DNA. Helicate identity is indicated on the absorbance figure for each row. The DNA base:helicate ratio for absorbance and ICD are indicated on the absorbance figures, that for the *LD* are shown on the *LD* figures. Absorbance and *CD* spectra were at 20 μM helicate with DNA concentrations to ensure the ratios indicated on the figure in the absorbance insert. A 1 cm path length cuvette was used, a water baseline was subtracted from each sample spectrum and the data zeroed at 420 nm. *LD* spectra were recorded at 500 μM DNA concentration and helicate concentrations to ensure the ratios indicated in the Figure. A 1 mm pathlength Couette flow cell[25] was used, a water baseline was subtracted from each sample spectrum and the data zeroed at 420 nm.

Fig 5 *ICD* spectrum of (R)-, (M)-, (P)- and [(M) + (P)]/2 for (a) L^3 and (b) L^5. ct-DNA = 200 μM, helicate = 20 μM. Spectra were collected in a 1 cm path length cell, a water baseline was subtracted from each sample spectrum and spectra were zeroed at 420 nm.

Fig 6 *ICD* spectra of (a) (R)-L^3, (b) (M)-L^3, (c) (P)-L^3, (d) (R)-L^5, (e) (M)-L^5 and (f) (P)-L^5 binding to ct-, AT- and GC-DNA at constant DNA (400 μM) and constant metal complex (20 μM) concentration. Spectra were collected in a 1 cm path length cuvette and a water spectrum was subtracted from each sample spectrum. The spectra were zeroed at 420 nm.

Fig 7 The structures of (a) Ethidium bromide and (b) Hoechst 33258.

Fig 8 Ethidium bromide fluorescence at 590 nm for different concentrations of (R)-, (M)- and (P)-L^1, L^3 and L^5 with constant DNA (12 μM) and EB (15 μM) concentration. Data were normalised to a starting intensity of 100.

Fig 9 Plot of fluorescence emission at the Hoechst maximum of 470 nm as a function of helicate concentration normalised to a starting intensity of 100 for comparison. 80 μM DNA, 8 μM Hoechst 33258. Spectra were collected in a 1 cm path length cuvette and the intrinsic Hoechst 33258 fluorescence spectrum subtracted from each sample spectrum.

Fig 10 *CD* spectra of GC (60 μM) in 50 mM NaCl and 1 mM cacodylate buffer with additions of $[Co(NH_3)_6]^{3+}$ (5 – 50 μM) with (a) no helicate, (b) (R)-L^3 (6 μM), and (c) (M)-L^5 (6 μM). Spectra were collected in a 1 cm masked cuvette and a water base line subtracted from each sample spectrum.

Fig 11 Plot of *CD* intensity at 290 nm as a function of $[Co(NH_3)_6]^{3+}$ concentration. The data were normalised to a starting intensity of 1 for comparison. CoH = $[Co(NH_3)_6]^{3+}$.

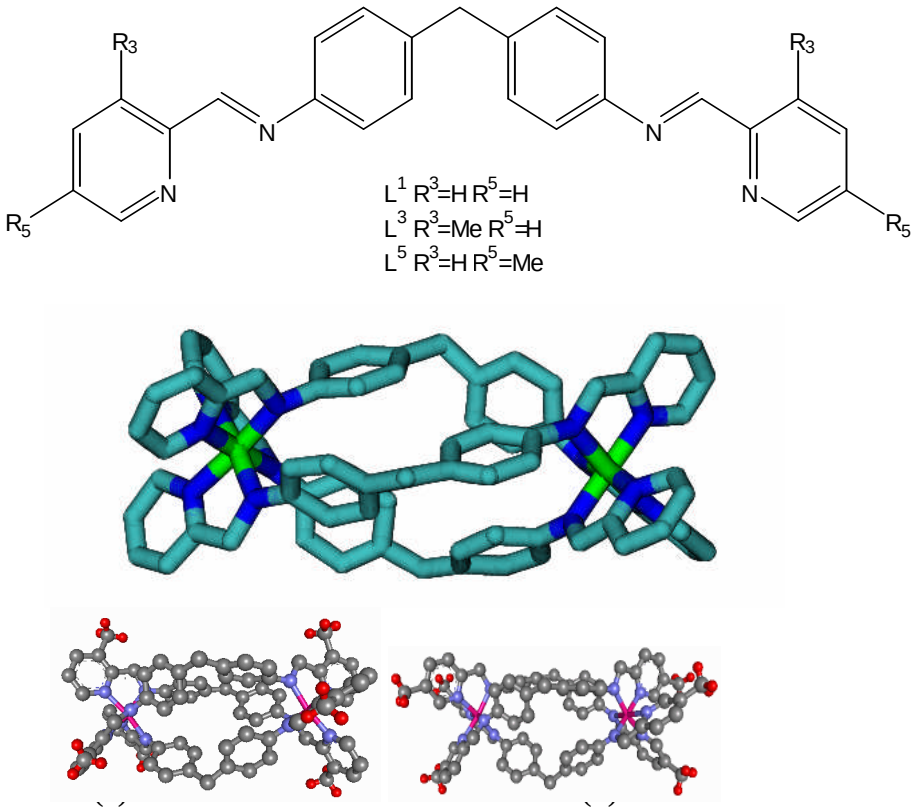


Fig 1

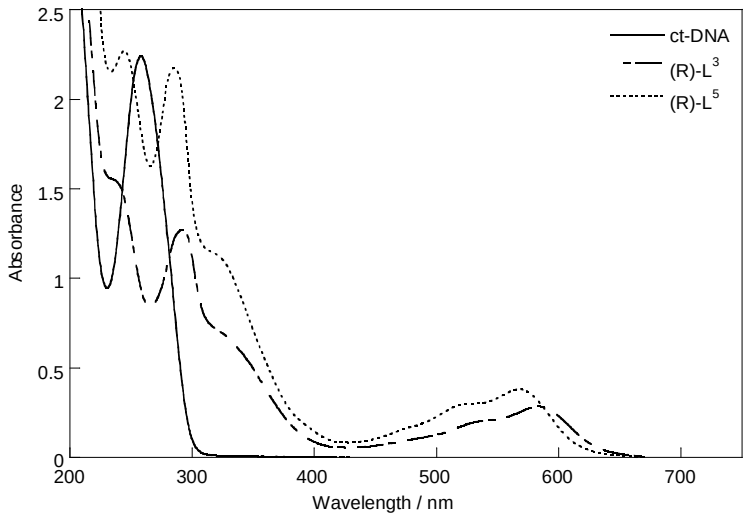

2.5
2
1.5
1
0.5
0
Absorbance
200
300
400
500
600
700
Wavelength / nm
ct-DNA
(R)-L3
(R)-L5

Fig 2

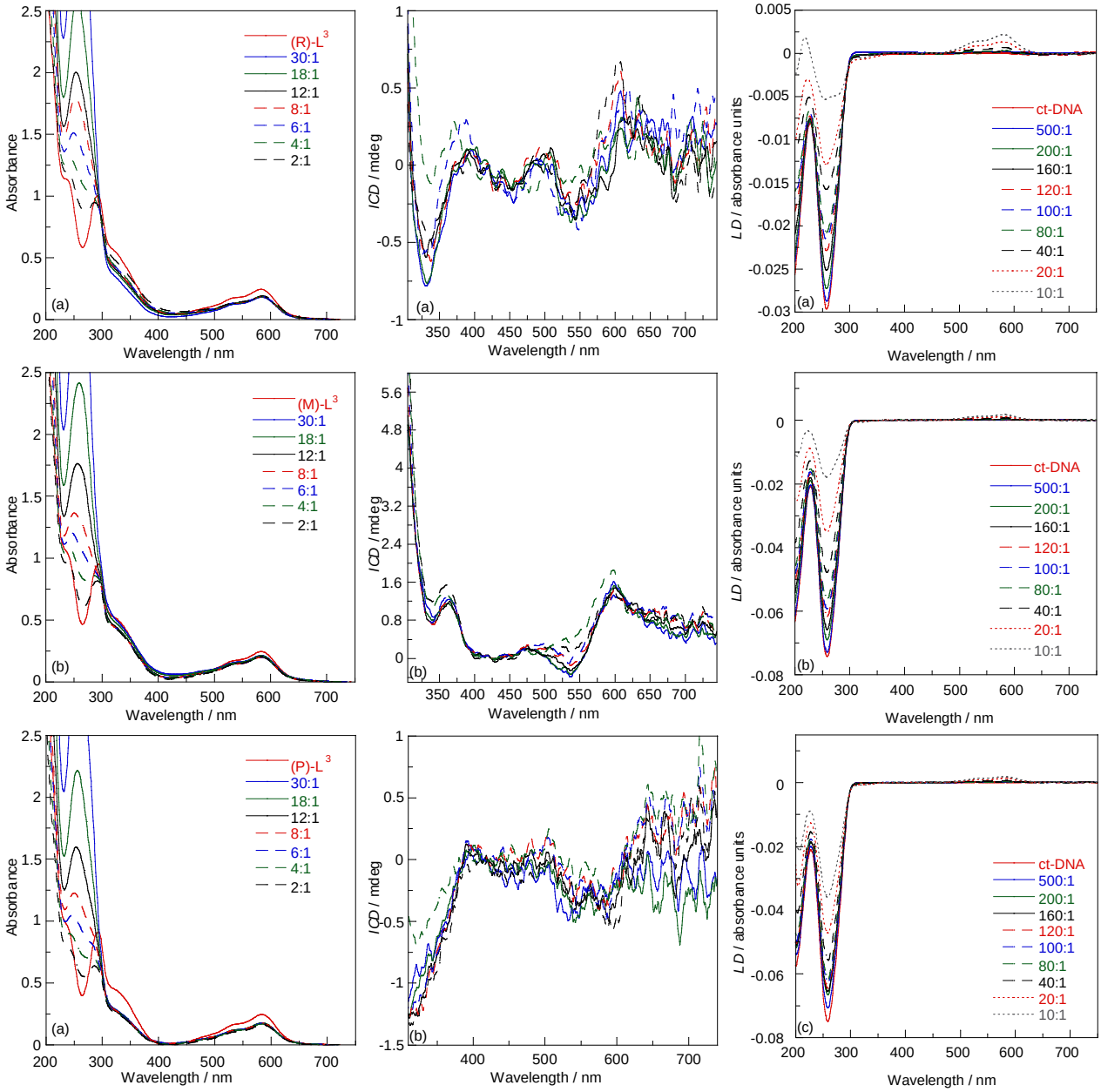

Absorbance
Wavelength / nm
(R)-L3
30:1
18:1
12:1
8:1
6:1
4:1
2:1
(a)
ICD / mdeg
LD / absorbance units
ct-DNA
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1
(M)-L3
(b)
(P)-L3
(c)

Fig 3

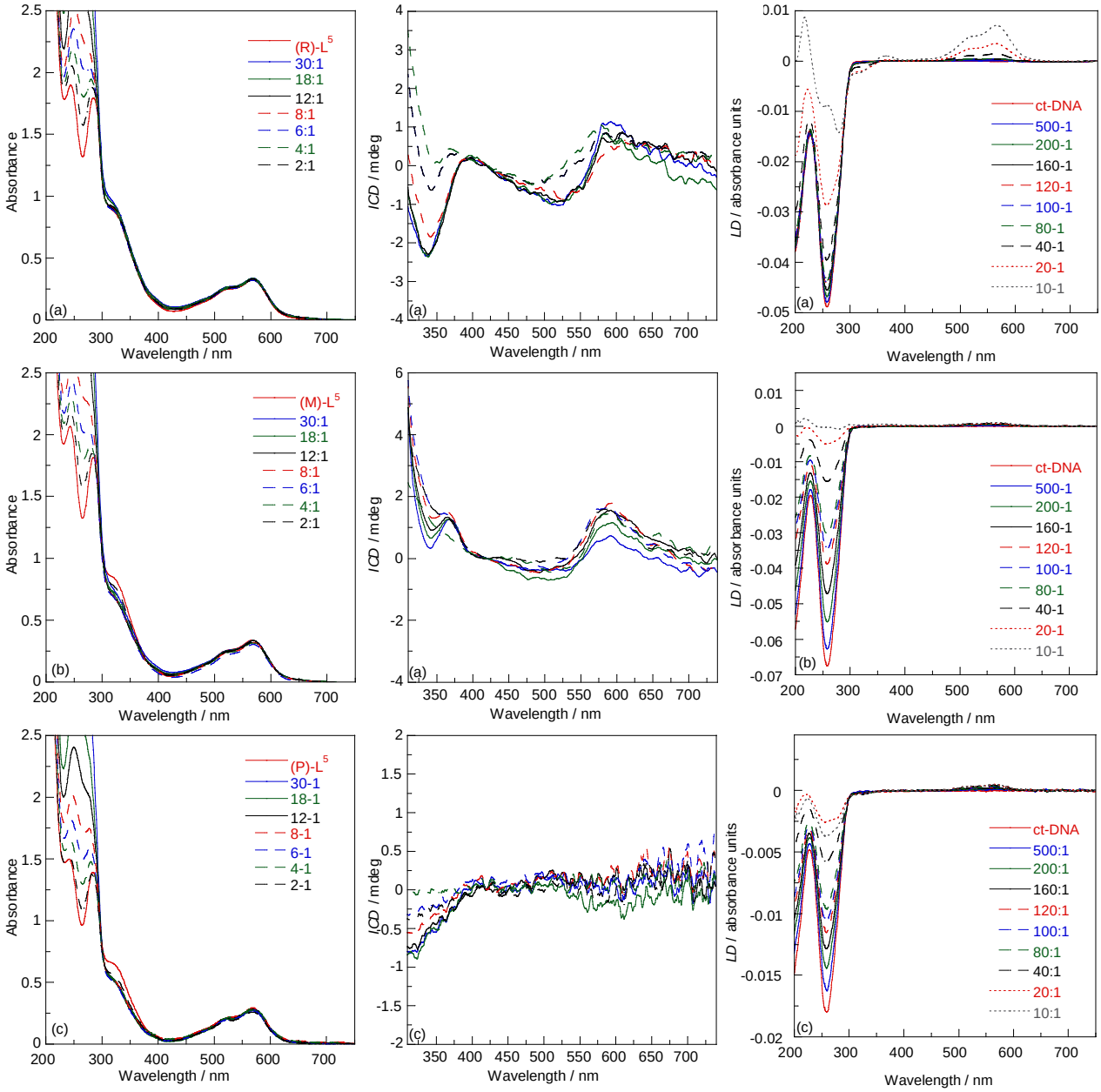

Absorbance
Wavelength / nm
(R)-L5
30:1
18:1
12:1
8:1
6:1
4:1
2:1
(a)
ICD / mdeg
LD / absorbance units
ct-DNA
500-1
200-1
160-1
120-1
100-1
80-1
40-1
20-1
10-1
(M)-L5
(b)
(P)-L5
30-1
18-1
12-1
8-1
6-1
4-1
2-1
(c)
500:1
200:1
160:1
120:1
100:1
80:1
40:1
20:1
10:1

Fig 4

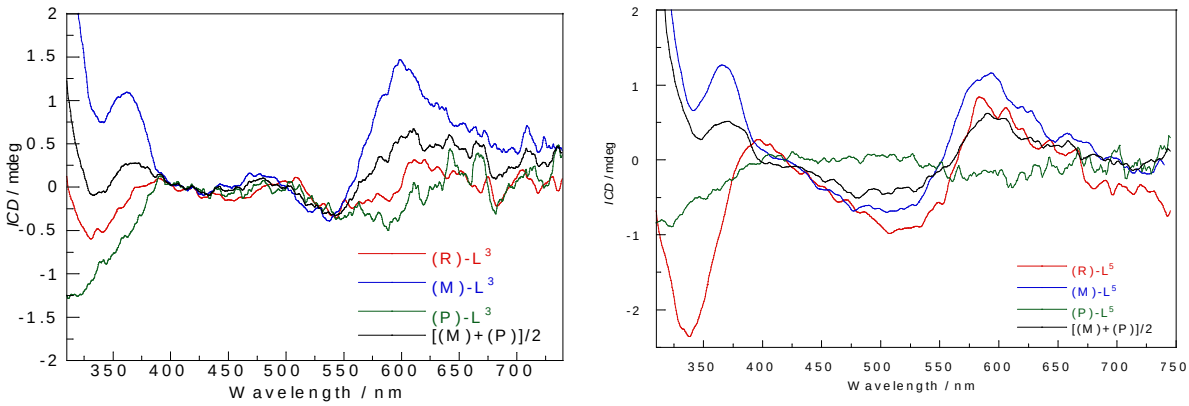


Fig 5

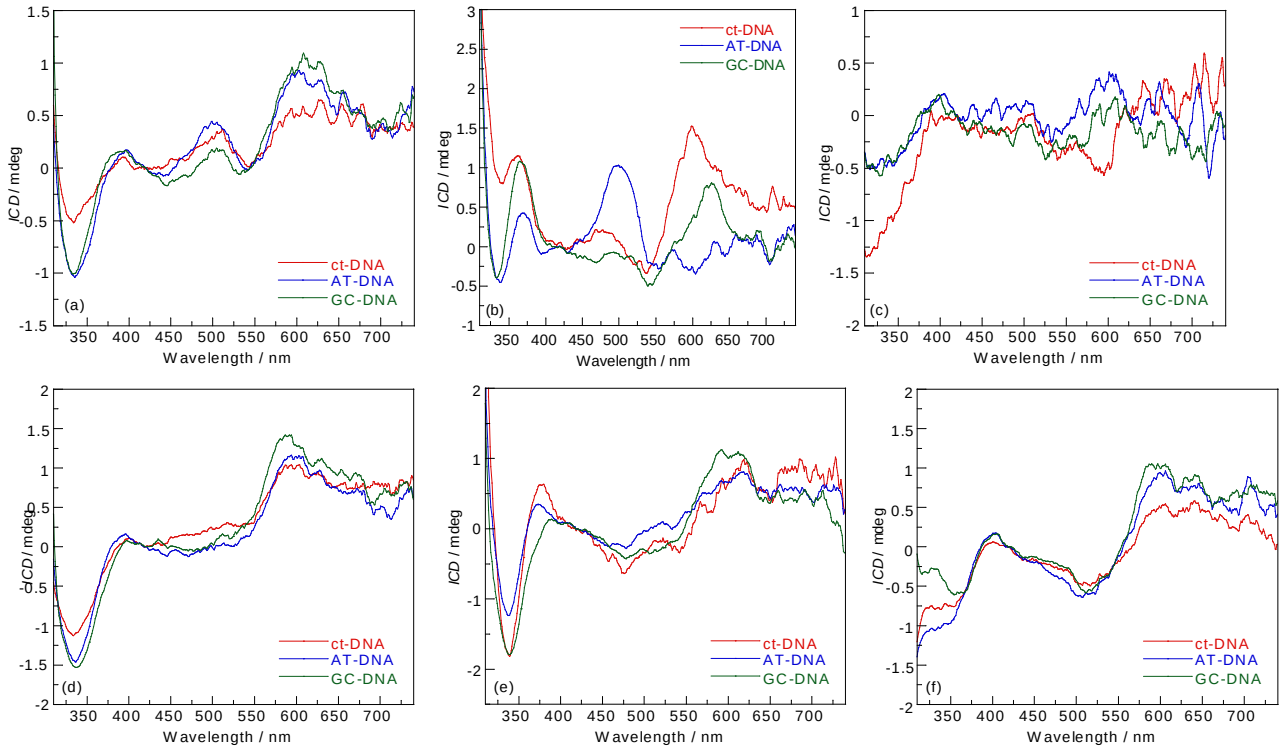
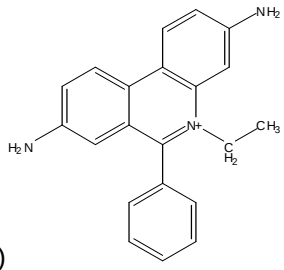
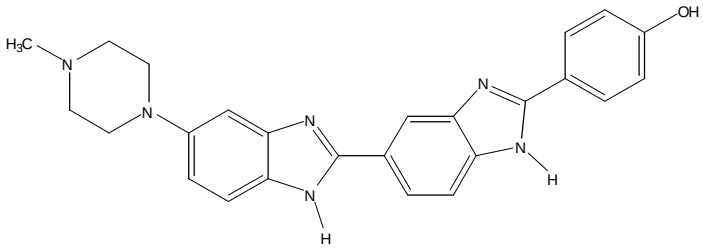


Fig 6



(a)



(b)

Fig 7

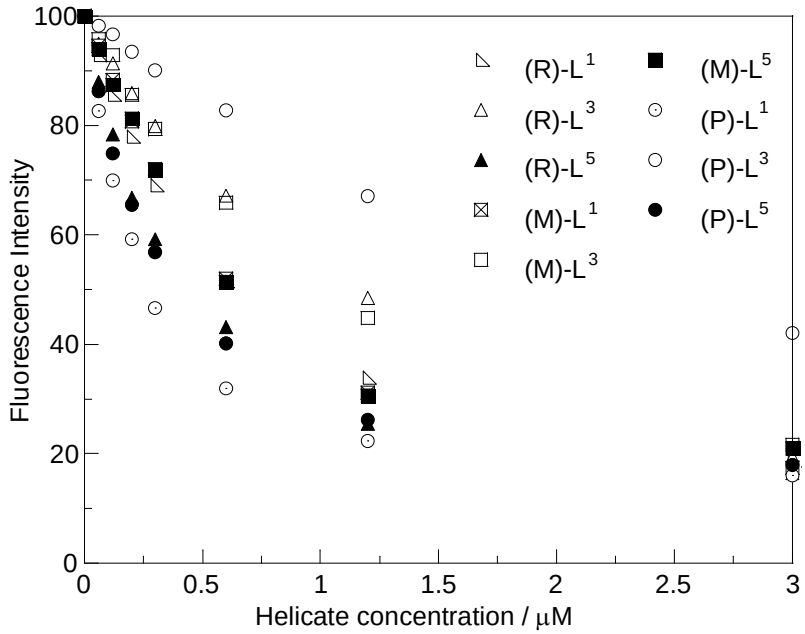

Fluorescence Intensity
Helicate concentration / μM
0
20
40
60
80
100
0
0.5
1
1.5
2
2.5
3
(R)-L1
(R)-L3
(R)-L5
(M)-L1
(M)-L3
(M)-L5
(P)-L1
(P)-L3
(P)-L5

Fig 8

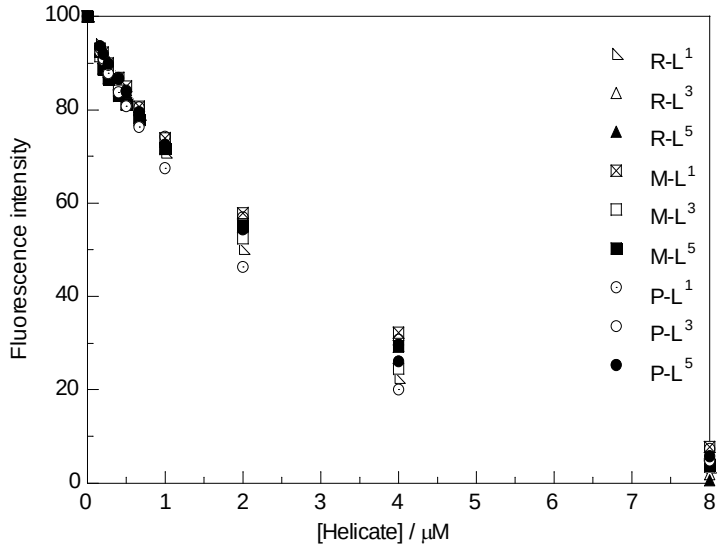

Fluorescence intensity
[Helicate] / μM
R-L1
R-L3
R-L5
M-L1
M-L3
M-L5
P-L1
P-L3
P-L5

Fig 9

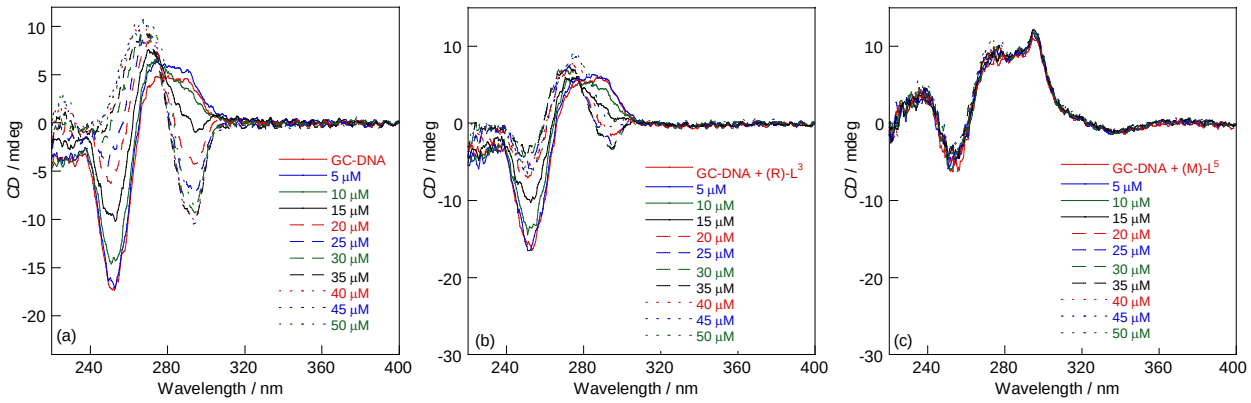

CD / mdeg
Wavelength / nm
(a)
GC-DNA
5 μM
10 μM
15 μM
20 μM
25 μM
30 μM
35 μM
40 μM
45 μM
50 μM
(b)
GC-DNA + (R)-L3
(c)
GC-DNA + (M)-L5

Fig 10

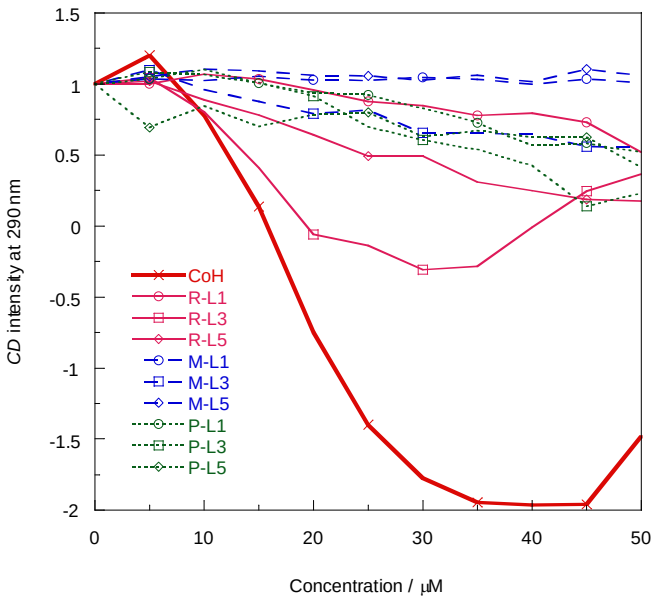

1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
10
20
30
40
50
CD intensity at 290 nm
Concentration / μM
CoH
R-L1
R-L3
R-L5
M-L1
M-L3
M-L5
P-L1
P-L3
P-L5

Fig 11

Publication 124

DNA Binding Studies of a New Dicationic Porphyrin. Insights into Interligand Interactions†

Alexander H. Shelton,‡ Alison Rodger,§ and David R. McMillin*,‡

Department of Chemistry, 560 Oval Drive, Purdue University, West Lafayette, Indiana 47907-2084, and Department of Biophysical Chemistry, MOAC Doctoral Training Centre, University of Warwick, Coventry CV4 7AL, U.K.

Received February 10, 2007; Revised Manuscript Received May 17, 2007

ABSTRACT: Cationic porphyrins have an affinity for DNA and potential for applications in the fields of photodynamic therapy and cellular imaging. This report describes a new dicationic porphyrin, 5,15-dimethyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrin, abbreviated H_2tMe_2D4. Although tetrasubstituted, H_2tMe_2-D4 presents modest steric requirements and forms in reasonable yield by a "2+2" synthetic method. Accordingly, studies of the zinc(II)- and copper(II)-containing derivatives, $Zn(tMe_2D4)$ and $Cu(tMe_2D4)$, have also been possible. Methods used to characterize DNA-binding motifs include absorption, emission, linear, and circular dichroism spectroscopies, as well as viscometry. An unusually detailed picture of porphyrin uptake emerges. As the ratio of DNA to porphyrin increases during a typical titration, H_2tMe_2-D4 or $Cu(tMe_2D4)$ initially aggregates on the host and then shifts to intercalative binding at close quarters before finally dispersing into non-interacting intercalation sites of the host. Emission studies of the copper(II) porphyrin have been very valuable. The existence of a measurable signal is diagnostic of intercalative binding, and the saturation behavior establishes that internalization typically monopolizes approximately three base pairs. In the moderate loading regime, emission data are most telling because dipole−dipole interactions between near-neighbor porphyrins tend to confuse other spectroscopic assays. The third ligand, $Zn(tMe_2D4)$, behaves differently in that the uptake is a strictly cooperative process. The mode of binding also varies with the base content of the DNA host. When the DNA is rich in A=T base pairs, the porphyrin remains five-coordinate and binds externally; however, $Zn(tMe_2D4)$ loses its axial ligand and binds by intercalation if the host contains only G≡C base pairs.

Water-soluble, cationic porphyrins are useful as chemical and spectroscopic probes of DNA structure (*1*−*4*). They also offer promise for various therapeutic applications, e.g., as sensitizers for photodynamic therapy (*5*, *6*), antibacterial agents (*7*), or inhibitors of teleomerase (*8*), an enzyme that plays a key role in extending the lifetime of tumor cells (*9*). Early work by Fiel and co-workers on 5,10,15,20-tetra(*N*-methylpyridinium-4-yl)porphyrin (H_2T4[1] in Scheme 1) stimulated many subsequent studies (*10*, *11*). As discussed in the reviews already cited (*1*−*4*), the consensus view is that H_2-T4 or one of its derivatives interacts with a particular B-form DNA duplex by one of three limiting binding motifs: intercalation between base pairs, external (groove) binding, or aggregation on the surface of the DNA macromolecule. Cooperative binding interactions can also occur. Many factors influence adduct formation, and one of the most important considerations is the composition of the DNA host. The empirical findings are that the presence of adenine·thymine (A=T) base pairs promotes external binding of H_2T4, while intercalation is more likely to occur within sequences that are rich in guanine·cytosine (G≡C) base pairs (*12*−*16*). The likely explanation is that the structural reorganization required for formation of a high-affinity external binding pocket is simply more feasible in a low-melting region of DNA, i.e., one that is rich in A=T base pairs (*17*−*19*). On the other hand, more rigid runs of DNA, which are rich in G≡C base pairs, tend to *disfavor* external binding and promote intercalative binding (*18*, *20*, *21*). The ionic strength of the solution is another important factor that influences binding. At lower ionic strengths, intercalative binding tends to be more favored (*22*, *23*), while external binding or even assembly of the porphyrin on the surface of the DNA surface becomes more favorable at higher ionic strengths (*24*).

Steric considerations are also important. The most obvious steric problems arise outside the plane of the porphyrin. Thus, axially ligated metalloporphyrins cannot intercalate between base pairs (*25*). Intercalation is apparently also impossible

† This work was supported by grants from the Engineering and Physical Sciences Research Council (GR/T09224/01 to A.R.) and the National Science Foundation (CHE-0550241 to D.R.M.).

* To whom correspondence should be addressed. E-mail: mcmillin@purdue.edu. Fax: (765) 494-0239. Phone: (765) 494-5452.

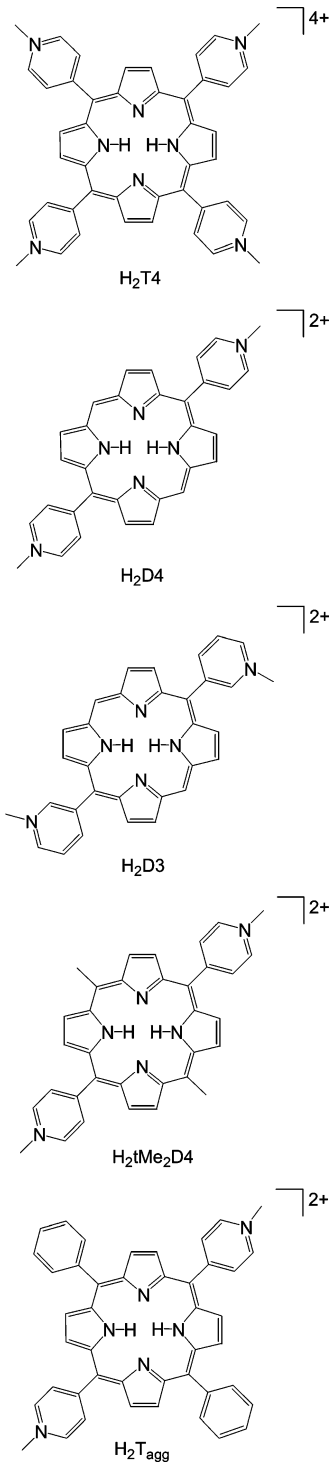
‡ Purdue University.

§ University of Warwick.

[1] Abbreviations: $H_2(tMe_2D4)$, 5,15-dimethyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrin; $Zn(tMe_2D4)$, [5,15-dimethyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrinato]zinc(II); $Cu(tMe_2D4)$, [5,15-dimethyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrinato]copper(II); CD, circular dichroism; H_2T4, 5,10,15,20-tetra(*N*-methylpyridinium-4-yl)porphyrin; CuT4, [5,10,15,20-tetra(*N*-methylpyridinium-4-yl)porphyrinato]copper(II); H_2D4, 5,15-di(*N*-methylpyridinium-4-yl)porphyrin; H_2D3, 5,15-di(*N*-methylpyridinium-3-yl)porphyrin; H_2T_{agg}, 5,15-diphenyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrin; DCM, dichloromethane; DMF, *N*,*N*-dimethylformamide; $CDCl_3$, deuterated chloroform; 5-Me-DPM, 5-methyldipyrromethane; H_2D4n, 5,15-di(4-pyridyl)porphyrin; H_2tMe_2-D4n, 5,15-dimethyl-10,20-di(4-pyridyl)porphyrin; LD, linear dichroism; LD^r, reduced linear dichroism; SRV, standard reduced viscosity ratio; *q*, base pair-to-porphyrin ratio; *R*, porphyrin-to-base pair ratio.

10.1021/bi700293g CCC: $37.00
Published on Web 00/00/0000 PAGE EST: 12

Scheme 1



for porphyrins carrying too many *N*-methylpyridinium-2-yl groups (*23*). Other steric interactions are also possible, and the sheer size of a substituent like the *N*-methylpyridinium ion, whether the linkage is via its C2, C3, or C4 position, has long been recognized as a likely influence on DNA binding interactions (*11*). Williams and co-workers provided some insight when they determined the crystal structure of a DNA duplex combined with Cu(T4), the copper(II)-containing derivative of H_2T4 (*26*). The structure showed that fraying occurred at the end of the helices, and the conclusion was that fraying occurs because significant steric clashes occur within the minor groove where *N*-methylpyridinium-4-yl groups of the intercalator impinge on the sugar−phosphate backbone of DNA. In a subsequent study, Guliaev and Leontis used ^{1}H NMR methods to determine the structure of an intercalated form of H_2T4 in solution; however, they did not report any particular steric problems (*27*). The latter study raises the possibility that the end fraying effect is a packing-induced artifact that occurs in the solid state. To address the issue experimentally, the McMillin group synthesized the sterically less demanding porphyrins 5,15-di(*N*-methylpyridinium-4-yl)porphyrin and 5,15-di(*N*-methylpyridinium-3-yl)porphyrin, denoted H_2D4 and H_2D3, respectively, in Scheme 1 (*18*, *28*). Note that the latter porphyrins need to insert only one *N*-methylpyridiniumyl substituent into the minor groove to intercalate into B-form DNA. Consistent with the idea that the steric bulk of H_2T4 inhibits intercalation in certain domains, H_2D4 and H_2D3 both act as universal intercalators (*18*, *28*). In other words, both porphyrins bind as intercalators, regardless of the base composition of the DNA host.

Unfortunately, progress in the area of sterically less demanding porphyrins has been slow. One reason is a materials limitation in that the synthetic yields reported for the H_2D4 and H_2D3 derivatives are very low. Studies are also time intensive because characterization of an adduct usually depends upon the application of an array of physical methods, most often absorption spectroscopy, circular dichroism (CD), and viscometry (*3*, *4*). Furthermore, the nature of the adduct may change with conditions, e.g., the extent of loading. The investigation reported below addresses these problems and describes a promising new porphyrin, 5,15-dimethyl-10,20-di(*N*-methylpyridinium-4-yl)porphryin, denoted herein as H_2tMe$_2$D4 (Scheme 1). The relative ease of preparation has made possible the extension of studies to include the copper(II) and zinc(II) derivatives, Cu(tMe$_2$D4) and Zn(tMe$_2$D4), respectively. The H_2tMe$_2$D4 system also makes an interesting contrast with 5,15-diphenyl-10,20-di(*N*-methylpyridinium-4-yl)porphyrin, or H_2T_{agg} in Scheme 1, a formally related porphyrin that is exceptionally prone to assembling on the DNA surface (*25*, *29*). The less hydrophobic H_2tMe$_2$D4 porphyrin is generally less prone to aggregation; however, Cu(tMe$_2$D4) has a tendency to aggregate in solution, and near-neighbor interactions occur for other H_2tMe$_2$D4 species under conditions of high drug loading. The luminescence properties of Cu(tMe$_2$D4) provide unique insight into adducts present in the high-loading regime where ligand−ligand interactions are inevitable. While H_2tMe$_2$D4 and Cu(tMe$_2$D4) behave as universal intercalators, the uptake of Zn(tMe$_2$D4) proves to be a base-dependent, strictly cooperative process.

EXPERIMENTAL PROCEDURES

Materials

Sigma-Aldrich sold 2-acetylpyrrole, 4-pyridinecarboxaldehyde, acetaldehyde, boron trifluoride diethyl etherate, copper acetate, dichlorodimethylsilane, florisil, methyl-*p*-toluenesulfonate, potassium hexafluorophosphate, sodium borohydride, tetrabutylammonium nitrate, Trizma base, Trizma HCl, and aluminum-backed silica gel thin layer chromatography plates with a fluorescent indicator. Pyrrole came from either Lancaster Synthesis or Acros Organics. Acetic anhydride, acetone, acetonitrile, ammonium hydrox-

ide, chloroform, dichloromethane (DCM), ether, ethyl acetate, glacial acetic acid, hexane, hydrochloric acid, methanol, *N*,*N*-dimethylformamide (DMF), potassium carbonate, potassium nitrate, silica gel, sodium carbonate, sodium hydroxide, sodium sulfate, and zinc acetate were products from Fisher or Mallinckrodt. Deuterated chloroform ($CDCl_3$) was from Sigma-Aldrich or Cambridge Isotope Labs. GE Healthcare (formally Amersham Biosciences) supplied sonicated salmon testes DNA (ST DNA) and poly(deoxyguanylic-deoxycytidylic) acid sodium salt {[poly(dG-dC)]$_2$}, while the poly(deoxyadenylic-deoxythymidylic) acid sodium salt {[poly(dA-dT)]$_2$} was from Midland Certified Reagent Company, Inc. The micron filters (0.22, 0.45, and 0.6 μm) were from Millipore Corp. or Pall Corp. Except as noted, commercial chemicals saw no purification. However, the deionized water did pass through a Barnstead Bantam water purification system with a mixed-bed cartridge.

Synthetic Procedures

Various literature reports helped shape the approaches taken for the synthesis of 2-(1-hydroxyethyl)pyrrole (*30*), 5-methyldipyrromethane (*30*), and 5,15-dimethyl-10,20-di(pyrid-4-yl)porphyrin (*18*, *23*, *31*, *32*).

2-(1-Hydroxyethyl)pyrrole. Rapid stirring induced 2-acetylpyrrole (25 g, 0.229 mol) to dissolve in a mixture of diethyl ether (229 mL) and water (678 mL). After the addition of sodium borohydride (34.652 g, 0.916 mol) in water (229 mL), the mixture was stirred for 48 h while reduction occurred. Subsequent steps involved extraction into ether and drying with sodium sulfate. After the drying agent had been filtered off, evaporation of the ether yielded the desired alcohol as an oil (84%): ^{1}H NMR ($CDCl_3$) δ 1.48 (d, 3H), 2.56 (s, 1H), 4.79 (m, 1H), 6.02 (s, 1H), 6.11 (q, 1H), 6.67 (d, 1H), 8.77 (s, 1H).

5-Methyldipyrromethane (5-Me-DPM). The reaction proceeded during dropwise (over the period of 1 h) addition of 9 g (0.081 mol) of 2-(1-hydroxyethyl)pyrrole to a flask containing 1.0 mL (8.1 mmol) of boron trifluoride diethyl etherate and 112.4 mL (1.62 mol) of pyrrole. The reaction time was 12 h with stirring under a nitrogen atmosphere. Subsequent workup involved the addition of 200 mL of dichloromethane and three washes with aqueous solutions containing 0.05 M sodium hydroxide and 10% (by mass) sodium carbonate. After drying over sodium sulfate, the organic layer yielded a dark viscous oil. After two extractions from florisil in conjunction with hexane and a Soxhlet extractor, a pale yellow solid slowly formed as the hexane evaporated (56% yield): ^{1}H NMR ($CDCl_3$) δ 1.70 (d, 3H), 4.30 (q, 1H), 6.17 (s, 2H), 6.25 (q, 2H), 6.76 (q, 2H), 7.96 (broad s, 1H).

5,15-Dimethyl-10,20-di(pyrid-4-yl)porphyrin (H_2tMe$_2$D4n). The reaction flask usually contained glacial acetic acid (300 mL), acetic anhydride (30 mL), and 4-pyridinecarboxaldehyde (0.42 mL, 4.2 mmol). The addition of the other component, 5-Me-DPM, occurred in portions and involved addition of aliquots of a 5 mL solution in glacial acetic acid prior to each of a series of heating steps. The heating times increased from 45 s to 2 min for reaction in a microwave oven under air. The cooling time was 30 min after each heating step. The workup steps include evaporation of the acid medium, dissolution in dichloromethane, filtration, and extraction with an aqueous base. Successive elutions from florisil with 2% methanol in DCM gave the desired porphyrin. The last purification step involved crystallization from an 8% dichloromethane/methanol solvent system. The final product deposited as a shiny purple amorphous solid (ca. 4% yield): calcd for $C_{32}H_{24}N_6 \cdot 1.5H_2O$ 73.98% C, 5.23% H, 16.17% N, found 74.19% C, 5.05% H, 16.00% N; ^{1}H NMR ($CDCl_3$) δ −2.64 (s, 2H), 4.64 (s, 6H), 8.14 (d, 4H), 8.82 (d, 4H), 9.05 (d, 4H), 9.51 (d, 4H); CI-MS ($C_{32}H_{24}N_6$) 492; $\epsilon_{416.5}$ (DCM) = 312 000 M^{-1} cm^{-1}.

[5,15-Dimethyl-10,20-di(N-methylpyridinium-4-yl)porphyrin](p-toluenylsulfonate)$_2$·H$_2$O. A modified literature method was used to methylate the neutral dipyridyl porphyrin (*18*). In addition to DMF as a solvent and methyl *p*-toluenesulfonate as the methylating agent, the reaction solution contained potassium carbonate. The reaction temperature was 60 °C under an atmosphere of nitrogen. It was convenient to follow the course of the reaction by a TLC method (*33*). After filtration, addition of a soluble nitrate salt induced precipitation of the product. Purification involved dissolution in water and treatment with dilute hydrochloric acid. Dropwise addition of aqueous potassium hexafluorophosphate induced precipitation of the red/brown hexafluorophosphate salt. Conversion to the nitrate salt was possible by a metathesis procedure: calcd for $C_{48}H_{44}N_6O_6S_2 \cdot H_2O$ 65.30% C, 5.25% H, 9.52% N, found 65.28% C, 5.42% H, 9.58% N; MALDI-MS ($C_{32}H_{24}N_6$) 522.2; ^{1}H NMR (DMSO-d_6) δ −3.62 (s, 2H), 3.88 (s, 6H), 3.92 (s, 6H), 8.17 (d, 4H), 8.20 (d, 4H), 8.65 (d, 4H), 9.11 (d, 4H).

{*[5,15-Dimethyl-10,20-(N-methylpyridinium-4-yl)porphyrinato]Zn(II)*}*(NO$_3$)$_2$·0.5H$_2$O*. The preparation involved the combination of H_2tMe$_2$D4 (59.9 mg, 92.7 μmol) with zinc acetate (148.0 mg, 927.0 μmol) in water and heating to 60 °C under nitrogen. After the reaction was over and the solution cooled to room temperature, the product deposited on addition of an aqueous solution of potassium hexafluorophosphate. A subsequent metathesis procedure then yielded the nitrate salt: calcd for $C_{34}H_{28}N_8O_6Zn \cdot 0.5H_2O$ 56.80% C, 4.06% H, 15.58% N, found 57.24% C, 4.13% H, 15.32% N.

{*[5,15-Dimethyl-10,20-(N-methylpyridinium-4-yl)porphyrinato]Cu(II)*}*(PF$_6$)$_2$·0.5H$_2$O*. The reaction was similar to that for the incorporation of zinc except that the solvent was DMF. Addition of tetrabutylammonium nitrate induced precipitation of the complex from DMF, and a metatheis procedure yielded the hexafluorophosphate salt: calcd for $C_{34}H_{28}N_6P_2F_{12}Cu \cdot 0.5H_2O$ 46.24% C, 3.31% H, 9.52% N, found 45.93% C, 3.38% H, 9.24% N.

Methods

Silanization of the cuvettes and other selected glassware helped minimize absorption of the cationic porphyrins during spectral studies (*34*). The extinction coefficients used for concentration determinations were as follows: ϵ_{260} = 13 600 M^{-1} cm^{-1} for [poly(dA-dT)]$_2$ (*35*), ϵ_{254} = 16 800 M^{-1} cm^{-1} for [poly(dG-dC)]$_2$ (*36*), and ϵ_{260} = 13 200 M^{-1} cm^{-1} for ST DNA (*37*), all in units of base pairs. Beer's law plots yielded the following for Zn(tMe$_2$D4) in Tris buffer (pH 7.50, μ = 0.05 M): $\epsilon_{416.5}$ = 190 000 M^{-1} cm^{-1} for H_2tMe$_2$D4, $\epsilon_{418.5}$ = 170 000 M^{-1} cm^{-1} for Cu(tMe$_2$D4), and ϵ_{428} = 150 000 M^{-1} cm^{-1}. The solvent used for determination of

the extinction coefficient for the copper analogue was 1:1 methanol/Tris buffer (pH 7.50, $\mu = 0.05$ M).

The buffer for the DNA binding studies was a $\mu = 0.05$ M, pH 7.50 Tris buffer. The porphyrin concentration was typically 1.7 μM for spectrophotometric investigations. The calculated %H, or hypochromic response, represents the percent drop in absorbance at the Soret maximum due to formation of an adduct with DNA. For quantifying all spectral *changes* associated with adduct formation, the reference was usually free porphyrin in Tris buffer. The exception was the hypochromism associated with the binding of Cu(tMe_2D4), in which case the reference solution was 50% MeOH, to ensure comparison with a monomeric form, vida infra. During luminescence measurements, the bandpass was 5 nm for both the excitation and emission slits in the experiments, except for studies involving the copper porphyrin when the slit was 20 nm on the excitation side. Equation 1 afforded corrections for the influence of absorbance

$$I_c = \frac{I}{1 - 10^{-A}} \quad (1)$$

where I is the experimental emission intensity, I_c is the corrected emission intensity, and A is the absorbance at the excitation wavelength.

For viscometry studies, solutions contained ST DNA at a concentration of 70 μM in base pairs. The mean length of the DNA molecules was ~3000 bp as established by gel electrophoresis by the supplier (GE Healthcare). Equation 2 gave the calculated standard reduced viscosity (SRV) ratio:

$$\frac{\eta}{\eta_0} = \frac{t_c - t_0}{t_D - t_0} \quad (2)$$

where t_0 is the flow time of the buffer, t_D is the flow time of DNA in buffer, and t_c is the flow time of the DNA solution containing porphyrin (*38*). For each measurement at 28.0 °C, the experimental flow time was the average of the first three consecutive runs that agreed with each other to within ±1 s. The buffer was a $\mu = 0.05$ M, pH 7.50 Tris solution, and immersion in a standard water bath provided a constant-temperature environment. The Purdue University Campus-Wide Mass Spectrometry Center took all mass spectrometry measurements, and D. Lee performed all microanalyses.

Linear dichroism experiments involved measuring the difference in the absorption of light polarized parallel and perpendicular to the flow direction as expressed in eq 3:

$$LD = A_{||} - A_{\perp} = A \times LD^r \quad (3)$$

where LD denotes the linear dichroism at wavelength λ (*39*, *41*). Equation 3 also defines the reduced linear dichroism signal (LD^r) which is the LD response divided by the isotropic absorbance A, at each wavelength.

Instrumentation

A Varian Cary 100 Bio UV−visible spectrophotometer, a Jasco V-550, or a Varian Cary 100 Scan UV−visible spectrophotometer provided absorbance data. The fluorescence spectrophotometer was a Varian Cary Elipse model, complete with an R3896 phototube, and the circular dichroism (CD) spectropolarimeter was a JASCO model J810. The NMR spectrometer was an Inova 300 MHz unit. A modified Cannon-Fenske model 25 viscometer was used. Other routine equipment used included a Corning model 430 pH meter and a Kenmore microwave oven. The LD instrument was a modified Jasco J-715 spectropolarimeter used in conjunction with a quartz Couette flow cell. One experiment involved the use of a microvolume Couette flow cell (*41*).

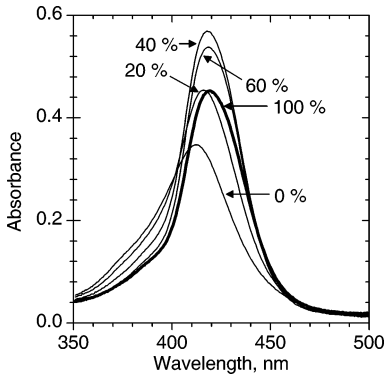


FIGURE 1: Absorption spectrum of a 4.0 × 10^{-6} M solution of Cu(tMe_2D4) as a function of volume percent of MeOH in 0.05 M Tris buffer (pH 7.5). The thick trace is 100% MeOH.

RESULTS

Porphyrin Synthesis. Combining 5-Me-DPM with 4-pyridinecarboxaldehyde to produce H_2tMe_2D4n is a variation of the 2+2 method for porphyrin synthesis (*30*, *42*, *43*) and is analogous to the route used previously for the synthesis of H_2D4n (*18*). The motivation behind the introduction of methyl substituents into positions 5 and 15 drew in part from a paper by Longo and co-workers (*44*), who proposed that tetrasubstituted derivatives like H_2tMe_2D4n naturally form in higher yield than a less substituted porphyrin such as H_2-D4. The argument is that, as condensation occurs, substituent groups induce coiling of the oligomer and thereby enhance the probability of ring closure. Although a subsequent paper cast doubt on the proposed substituent effect (*45*), the isolated yield of H_2tMe_2D4n is ~4% and consistently greater than (ca. 4 times) the yield obtained for H_2D4n. Methylation of H_2tMe_2D4n with methyl *p*-toluenesulfonate gives the dication H_2tMe_2D4 which exhibits a Soret absorption maximum at 416 nm in aqueous solution. The overall yield is modest but practicable for the synthesis and characterization of new metalated forms. As an aside, it is worth noting that Gonçalves et al. reported a higher-yield synthesis of H_2D4 after the current study was complete (*46*).

Heating a DMF solution containing H_2tMe_2D4 and excess zinc(II) or copper(II) acetate leads to insertion of zinc(II) or copper(II) into the porphyrin (*47*). Treating the free porphyrin with aqueous acid (*48*) before inserting copper minimizes the chances of contamination with the much stronger lumaphor Zn(tMe_2D4). The zinc derivative Zn(tMe_2D4) has its Soret maximum at 428 nm. The Soret band of the Cu(tMe_2D4) system is different in that it is comparatively broad and has a low apparent molar absorptivity in aqueous solution. The molar absorptivity of Cu(tMe_2D4) also increases in methanol (Figure 1), whereas the absorbance of H_2tMe_2D4 or Zn(tMe_2D4) decreases by ~3%. As Figure 1 shows, the absorption intensity of Cu(tMe_2D4) is actually

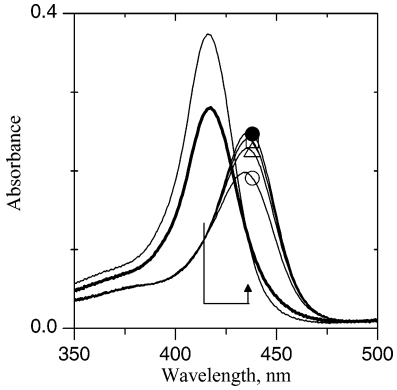


FIGURE 2: Absorbance of H_2tMe_2D4 in 0.05 M Tris buffer (pH 7.5) containing ST DNA at base pair-to-porphyrin ratios of 0 (thin trace), 1 (thick trace), 10 (○), 25 (△), 50 (□), and 100 (●).

highest in a solution that is ~40% methanol by volume. The absorbance increase at low methanol percentages is attributable to the dissociation of aggregated forms of Cu(tMe_2D4); see below for more evidence of aggregation phenomena. At higher methanol percentages, the absorbance decreases because the absorptivity of the monomer is inherently smaller in alcohol. In the absence of DNA, a 50% methanol/Tris buffer solution is a convenient solvent for absorbance comparisons with DNA-containing samples because Cu(tMe_2D4) exists as a monomer. The Soret maximum then occurs at 419 nm.

DNA Binding of H_2tMe_2D4. The data in Figure 2 reveal that the addition of ST DNA gives rise to three distinct modes of interaction with H_2tMe_2D4. The first occurs when the porphyrin is in excess and the base pairs-to-porphyrin quotient (q) is less than or equal to 1. Under these conditions, interaction with the host produces a hypochromic response in the Soret region but little or no spectral shift. There is also effectively no induced CD signal in the Soret region. The second type of adduct develops under conditions of intermediate loading ($2 \leq q \leq 10$), as evidenced both by a shift of the Soret band to a longer wavelength and the enhanced hypochromic response. Finally, the third mode of binding becomes evident when $q \geq 10$, when excess DNA is present in solution. In this regime, the wavelength of the Soret maximum shifts to a slightly longer wavelength, while the absorption intensity trends upward and asymptotically approaches a limiting hypochromicity (H) value of 32%. At the same time, the negative induced CD signal achieves a limiting amplitude ($\Delta\epsilon$) of -17 M^{-1} cm^{-1} at 440 nm (Table 1). Farther into the UV region, the induced CD spectrum includes a positive band at 370 nm. Binding also affects the emission of the porphyrin. The corrected emission spectrum of the free porphyrin is broad and weakly structured with a 0−0 band at 670 nm and a shoulder, with an intensity approximately half as great, at ~720 nm. In the presence of excess ST DNA, the emission intensity approximately doubles, and the resolution improves. The 0−0 band of the adduct occurs at 681 nm, while the adjoining 1−0 band appears at 746 nm and seems to weaken in intensity.

A similar sequence of events occurs with [poly(dG-dC)]$_2$ or [poly(dA-dT)]$_2$ as the host. The spectral changes are larger with the [poly(dG-dC)]$_2$ host, and by the time $q = 10$, the bathochromic shift of the Soret band is greater than 20 nm. With the addition of more [poly(dG-dC)]$_2$, the absorption maximum shifts further toward the red, while the absorption

Table 1: Spectral Data Obtained in 0.05 M Tris Buffer (pH 7.5)[a]

	$\Delta\lambda$ (nm) (%H)[b]	λ_{CD} (nm) [$\Delta\epsilon$ (M^{-1} cm^{-1})]	λ_{em} (nm)[c]
H_2tMe_2D4			670, 720sh
[poly(dA-dT)]$_2$	18 (14)	438 (−13)	679, 744
[poly(dG-dC)]$_2$	27 (40)	444 (−20)	690, 757
ST DNA	21 (32)	440 (−17)	681, 746
Cu(tMe_2D4)			
[poly(dA-dT)]$_2$	16 (32)	434 (−18)	821 broad
[poly(dG-dC)]$_2$	23 (52)	442 (−20)	822 broad
ST DNA	18 (50)	434 (−17)	820 broad
Zn(tMe_2D4)			670, 710sh
[poly(dA-dT)]$_2$	5 (15)	425 (+31) 449 (−25)	650, 705sh
[poly(dG-dC)]$_2$	18 (48)	~426 (+4) 454 (−5)	675, 719sh
ST DNA	9 (32)	424 (+6) 450 (−12)	657, 713sh

[a] When DNA is present, the data are for the limiting adducts formed at high base pair-to-porphyrin ratios. [b] Hypochromism (%H) is the percent change in absorbance at the absorption maximum. [c] From corrected spectra, except for the Cu(tMe_2D4) system which emits at longer wavelengths where the instrumental correction factors are less reliable.

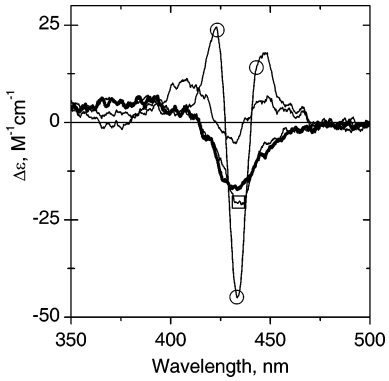


FIGURE 3: Induced CD spectra of Cu(tMe_2D4) in 0.05 M Tris buffer (pH 7.5) containing [poly(dA-dT)]$_2$ at q values of 1 (thin trace), 3 (○), 25 (□), and 100 (thick trace).

intensity increases and levels off when $H = 40\%$ by $q = 75$. Here, too, the limiting adduct shows two induced CD bands in the near UV, a negative band at 444 nm ($\Delta\epsilon = -20$ M^{-1} cm^{-1}) and a positive band at 370 nm ($\Delta\epsilon = 12$ M^{-1} cm^{-1}). Formation of an adduct with [poly(dA-dT)]$_2$ follows a similar pattern. At $q = 10$, the induced CD signal is multisignate, but the structure disappears when higher levels of [poly(dA-dT)]$_2$ are present. The limiting induced CD spectrum parallels that obtained with ST DNA or [poly(dG-dC)]$_2$ in exhibiting a negative band in the Soret region at 438 nm ($\Delta\epsilon = -13$ M^{-1} cm^{-1}) as well as a second band at a higher energy.

DNA Binding of Cu(tMe_2D4). Like H_2tMe_2D4, Cu(tMe_2D4) experiences three distinct binding environments during a DNA titration. At low levels of [poly(dA-dT)]$_2$, when $q < 2$, the induced CD signal is weak and trisignate, but it becomes much more intense during the intermediate phase of binding [$2 \leq q \leq 10$ (Figure 3)]. In the final phase of binding, by $q = 25$, the signal devolves into a negative band with the extreme at ~434 nm. In contrast, the emission signal exhibits a relatively simple progression. Prior to the addition of [poly(dA-dT)]$_2$, there is no detectable emission from Cu(tMe_2D4), as one would expect because of solvent-induced quenching (*4*, *49*, *50*). However, the signal begins to grow with the addition of [poly(dA-dT)]$_2$; see Figure 4. By $q =$

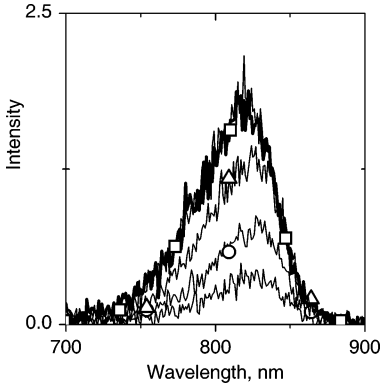


FIGURE 4: Absorbance-corrected emission spectra of Cu(tMe$_2$D4) in the presence of [poly(dA-dT)]$_2$ and 0.05 M Tris buffer (pH 7.5) at q values of 0.5 (thin trace), 1 (○), 2 (△), 10 (□), and 75 (thick trace). The $q = 10$ and 75 spectra virtually coincide.

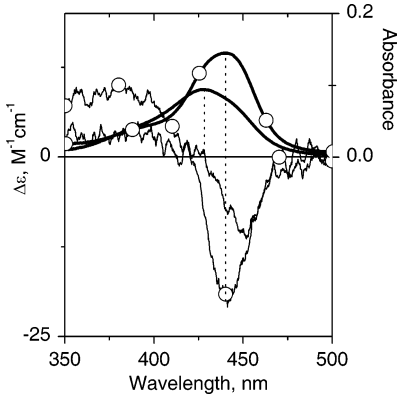


FIGURE 5: Induced CD spectra (thin traces) and absorbance spectra (thick traces) of Cu(tMe$_2$D4) in 0.05 M Tris buffer (pH 7.5) containing [poly(dG-dC)]$_2$ at q values of 2 (thick trace) and 75 (○). A dashed line relates corresponding spectra at the wavelength of the absorption maximum.

4, the signal essentially reaches a limiting emission intensity, albeit orders of magnitude weaker than the signal from H$_2$-tMe$_2$D4. The results differ only in detail when [poly(dG-dC)]$_2$ acts as the host. Most importantly, there is never an indication of a trisignate-induced CD signal. Instead, under conditions of intermediate loading ($2 \leq q \leq 10$), interaction with [poly(dG-dC)]$_2$ induces a bisignate but largely negative CD signal in the Soret region. The bisignate character is evident from the fact that the minimum occurs at around 450 nm, whereas the absorption is maximized at a significantly shorter wavelength (Figure 5). However, upon addition of excess DNA, the CD signal intensifies and shifts toward a higher energy until the band positions agree in the CD and absorption spectra. With this host, the emission signal from Cu(tMe$_2$D4) is maximized by $q \simeq 2$ and is ~3 times as intense as that obtained with [poly(dA-dT)]$_2$.

Qualitatively similar results are obtained with sonicated ST DNA as the host. Thus, the largest hypochromic response ($H \approx 67\%$) occurs in the absorption spectrum of Cu(tMe$_2$-D4) under high loading conditions ($q \leq 3$), when there is almost no shift in the wavelength of the Soret maximum. Even though the emission signal from Cu(tMe$_2$D4) stops changing by $q = 3$, the largest shift of the Soret absorption occurs in the intermediate-loading regime ($3 \leq q \leq 10$). In the intermediate-loading regime, the system also maintains an isosbestic point at 410 nm; however, a new isosbestic point develops at 419 nm when the base pair-to-copper ratio

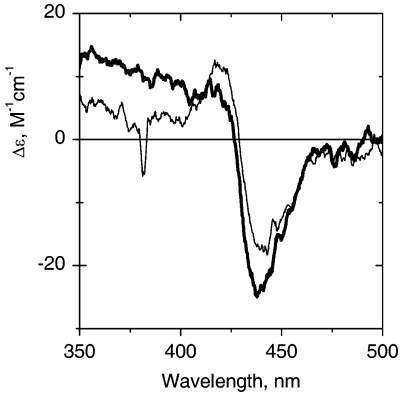


FIGURE 6: Induced CD spectrum of Cu(tMe$_2$D4) in 0.05 M Tris buffer (pH 7.5) containing ST DNA at a q value of 14 at ionic strengths of 0.05 M (thick trace) and 0.45 M (thin trace).

increases beyond 25. During formation of the final adduct, the hypochromism approaches a limiting value of ca. 50%. The CD results track those obtained with [poly(dG-dC)]$_2$ as the host.

Effect of Ionic Strength. Increasing the ionic strength has a modest effect on the spectral properties of Cu(tMe$_2$D4) interacting with ST DNA. As the ionic strength ranges from 0.05 to 0.45 M at $q = 14$, the Soret maximum shifts from 431 to 427 nm and the absorbance drops by 22%. Figure 6 portrays CD data obtained at either end of the ionic strength range. Bisignate character is apparent at $\mu = 0.45$ M where the spectrum shows a negative band at 442 nm as well as a weaker positive peak that is maximal around 420 nm. However, there is no shift of either the emission maximum or the emission intensity, after correcting for the absorbance change.

DNA Binding of Zn(tMe$_2$D4). Zn(tMe$_2$D4) is different in that it apparently exhibits only two phases of binding. When the porphyrin binds to [poly(dG-dC)]$_2$, there is no significant change in the emission spectrum, but there are changes in the electronic absorption. Thus, by $q \approx 10$, the Soret band experiences a bathochromic shift, and the hypochromic response sets in at even lower ratios. Compared with that of Cu(tMe$_2$D4), relatively high levels of [poly(dG-dC)]$_2$ are necessary to force formation of the limiting spectrum ($q \geq 100$). Nevertheless, binding induces sizable shifts in the absorbance maximum and intensity (Table 1). At saturation, the induced CD spectrum exhibits a minimum at ca. 454nm and crosses zero at around 443 nm, essentially at the wavelength of the absorption maximum. Thus, the absorbance and CD data obtained for the final adduct mimic those of the adduct that Cu(tMe$_2$D4) forms with the same [poly(dG-dC)]$_2$ host during the intermediate phase of binding. Zn(tMe$_2$D4) also experiences only two binding environments in a titration with [poly(dA-dT)]$_2$. However, the absorption spectrum exhibits a relatively small bathochromic shift and practically no hypochromism when the porphyrin binds to [poly(dA-dT)]$_2$. Binding to [poly(dA-dT)]$_2$ is also distinctive in that the emission signal shifts toward a shorter wavelength. Furthermore, all spectral changes are complete by $q \approx 50$. See Figure 7 for representative CD spectra. At saturation, the signal exhibits a positive shoulder at ~400 nm beside a moderately intense bisignate signal that crosses over the baseline at 437 nm, ~5 nm beyond the absorption maximum.

Combined Uptake. Simultaneously exposing DNA to Cu-(tMe$_2$D4) and Zn(tMe$_2$D4) provides a means of exploring

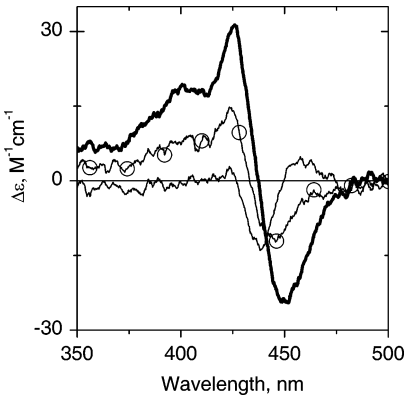


FIGURE 7: Induced CD spectra of Zn(tMe$_2$D4) in 0.05 M Tris buffer (pH 7.5) containing [poly(dA-dT)]$_2$ at q values of 1 (thin trace), 7 (○), and 50 (thick trace).

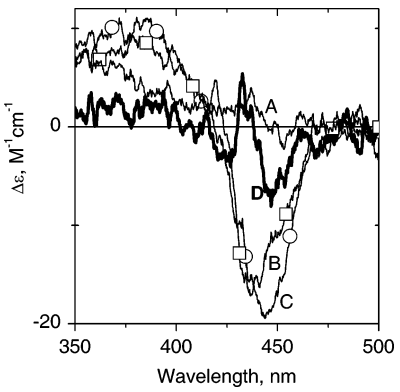


FIGURE 8: Induced CD spectra of porphyrins in 0.05 M Tris buffer (pH 7.5) with [poly(dG-dC)]$_2$ as the host. Samples contained Zn(tMe$_2$D4) at $q = 10$ (A, thin trace), Cu(tMe$_2$D4) at $q = 10$ (B, □), or Zn(tMe$_2$D4) and Cu(tMe$_2$D4), in which case the base pair-to-porphyrin ratio is 10 for each porphyrin (C, ○). The calculated difference spectrum (D, thick trace) represents spectrum C minus spectrum B.

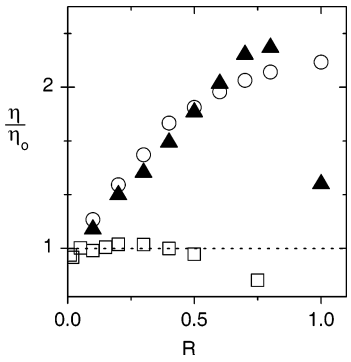


FIGURE 9: Standard reduced viscosity ratios as a function of the porphyrin-to-DNA base pair ratio R. The medium is 0.05 M Tris buffer (pH 7.5), and H$_2$tMe$_2$D4 (○), Cu(tMe$_2$D4) (▲), and Zn(tMe$_2$D4) (□) were employed as ligands.

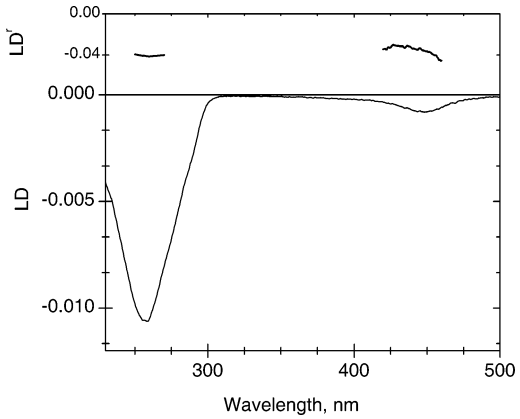


FIGURE 10: Linear dichroism of Zn(tMe$_2$D4) interacting with [poly(dG-dC)]$_2$ in 0.05 M Tris buffer (pH 7.8): (bottom) LD spectrum at a base pair-to-porphyrin ratio of 60 and (top) LD^r data for the same sample.

cooperative effects. No unusual effects are evident when [poly(dA-dT)]$_2$ acts as the host, but the shift in the fluorescence maximum suggests that the uptake of Zn(tMe$_2$D4) is virtually complete when the base pairs-to-porphyrin ratio is 10 for each porphyrin. There is definitely a specific interaction with Zn(tMe$_2$D4) because the DNA absorption at 260 nm appears in the fluorescence excitation spectrum of the porphyrin. Cooperative binding is easier to recognize when Zn(tMe$_2$D4) interacts with [poly(dG-dC)]$_2$, because the affinity is comparatively low. Here, CD data clearly reveal that uptake of the copper porphyrin promotes binding of Zn(tMe$_2$D4). Results in Figure 8 show, for example, that Zn(tMe$_2$D4) contributes to the CD signal when [poly(dG-dC)]$_2$ is the host and each porphyrin is present at a DNA base pair-to-porphyrin ratio of 10. However, there is no significant signal from Zn(tMe$_2$D4) at the same loading level in the absence of Cu(tMe$_2$D4).

Viscometry. Viscometry measurements are useful because intercalation of a ligand into the DNA macromolecule has a measurable effect on the flow properties (*51*, *52*). In this study, the solutions contain a fixed concentration of sonicated ST DNA, while the porphyrin-to-base pair ratio, R, varies. The results in Figure 9 show that the specific viscosity increases by a factor of ~2 with the addition of H$_2$tMe$_2$D4, or Cu(tMe$_2$D4), before leveling of at ca. $R = 0.8$. At higher loadings ($R = 1.0$), the viscosity drops sharply in the case of Cu(tMe$_2$D4), presumably because the porphyrin aggregates on the surface of the DNA. On the other hand, uptake of Zn(tMe$_2$D4) produces quite different results. Thus, there is only a hint of an increase in the specific viscosity at low R values, before a noticeable decrease begins to occur at $R >$ 0.4.

Linear Dichroism. Figure 10 shows LD and LD^r spectra of Zn(tMe$_2$D4) interacting with [poly(dG-dC)]$_2$ at $q = 60$. Results obtained with Zn(T4) and H$_2$T4 are very similar. Due to the flow-induced ordering of the polymer, slightly more DNA molecules on average have a parallel orientation in solution. As the $\pi-\pi^*$ transitions of the bases are in-plane polarized and the base pairs are all more or less perpendicular to the long axis of B-form DNA, the LD and LD^r signals in Figure 10 are negative in the vicinity of 260 nm, in line with eq 3 ($A_\perp > A_{||}$). The corresponding LD^r signal at 440 nm is due only to the porphyrin and is almost equally negative. It follows that the porphyin ligand is roughly perpendicular to the long axis of DNA, too, because the Soret band is also $\pi-\pi^*$ absorption of a planar, aromatic ring system. Note that the LD^r signal in Figure 10 actually underestimates the dichroic response in the Soret region because there is still free Zn(tMe$_2$D4) in solution at $q = 60$. (Other data obtained with a microvolume Couette cell revealed that the magnitude of the LD response in the Soret region signal increased by ca. 40% relative to that observed in the 260 nm region at $q = 75$. However, the latter experiment required a higher porphyrin concentration, and

a precipitate formed before it was possible to gather the absorbance data needed for calculation of an accurate LD^r spectrum.) The conclusion is that the LD data obtained at $\mu = 0.05$ M for Zn(tMe_2D4), Zn(T4), and H_2T4 are all consistent with intercalative binding to [poly(dG-dC)]$_2$.

DISCUSSION

The following questions remain: What modes of binding unfold as DNA takes up H_2tMe_2D4 or Cu(tMe_2D4)? Why is the uptake of Zn(tMe_2D4) a qualitatively different process? Do the steric demands of the methyl substituents at positions 5 and 15 of H_2tMe_2D4 impact the binding to DNA? And, is there any evidence of cooperative uptake? As a prelude to the resolution of these issues, a brief summary of some important background is useful.

Aggregation Effects

Aggregation is an important effect to consider in aqueous media because porphyrins have such a hydrophobic core, although the charged periphery provides a countervailing force. Sterically active substituents may also affect the process as indicated by the work of Tjahjono and co-workers (*53*), who found that the dication 5,15-bis(1,3-dimethylimidazolium-2-yl)porphyrin exists as a monomer in water whereas the more planar 5,15-bis(1,2-dimethylpyrazolium-4-yl)porphyrin undergoes extensive aggregation. In the case of the tetracationic porphyrin H_2T4, the repulsive effects completely dominate, because H_2T4 exists as a monomer under normal conditions (*54*, *55*). However, the dicationic analogue H_2T_{agg} behaves differently. It shows a strong tendency to aggregate as evidenced by a broadened Soret band, diminished absorbance intensity, and increased dynamic light scattering (*29*). Furthermore, codissolved DNA macromolecules enhance the effect by serving as templates for aggregation (*24*, *56*). Since H_2tMe_2D4 and its various derivatives are dicationic porphyrins, aggregation is something to consider.

Luminescence and Copper(II) Porphyrins

Luminescence studies of copper(II) porphyrins provide unique insight into DNA binding interactions. One reason is that the emission signal is extremely sensitive to the local environment about the copper center. Another plus, which will be important in the discussion below, is that through-space coupling with other chromophores does not affect the signal. The steric demands of the copper(II) form are also similar to those of the free porphyrin. Thus, in the electronic ground state, copper(II) porphyrins tend not to bind axial ligands because the "hole" in the d^9 shell of the copper center resides in the plane of the porphyrin, i.e., in the $d_{x^2-y^2}$ orbital. Thermal population of a five-coordinate structure is normally not feasible except in the presence of a very strong base such as pyridine (*57*). However, it is important to recognize that axial ligands have a profound influence on the excited state dynamics. The luminescence spectra of a free porphyrin and its copper(II) derivative are always very different. While the free porphyrin exhibits fluorescence from a $^1\pi-\pi^*$ excited state, the four-coordinate copper(II) derivative emits from a lower-energy state (with a multiplet structure) derived from the corresponding triplet intraligand excited state (*58*). Multiplicity changes are rapid in the copper(II) derivative due to the presence of the unpaired electron on the central metal ion. In addition, efficient quenching by Lewis bases, including water, is unique to the copper(II) form. Formation of a five-coordinate adduct is feasible in the excited state because the hole can migrate to the d_{z^2} orbital (*16*, *49*, *50*, *59*−*62*). The five-coordinate form is nonemissive because radiationless decay is facile and because d−d states have intrinsically low dipole strengths. Observation of the five-coordinate d−d excited state has been possible with a time-resolved X-ray absorption technique (*63*), as well as by time-resolved resonance Raman spectroscopy (*64*). Indeed, Raman studies have shown that formation of a photoexcited, five-coordinate form is virtually obligatory whenever a copper(II) porphyrin binds externally to single- or double-stranded DNA (*50*, *65*, *66*). The only way the emission survives is if the DNA host shields the axial positions of the copper(II) center from attack by solvent, buffer, and/or Lewis bases embedded in the DNA host. Observation of emission from bound Cu(T4) therefore proves to be the signature of intercalative binding (*16*, *49*, *62*). Szalai and co-workers (*67*) have used a similar strategy to establish that Cu(T4) is also capable of internalizing into quadruplex-forming host structures.

Interactions of DNA with H_2tMe_2D4 and Cu(tMe_2D4)

With all three porphyrins and H_2tMe_2D4 and Cu(tMe_2D4) in particular, the nature of the adduct that forms and the consequent spectral changes depend very much on the DNA base pair-to-porphyrin ratio (q) in solution. The interpretation is most straightforward when DNA is present in large excess.

Excess DNA Regime ($q \geq 25$). When excess DNA is present, both H_2tMe_2D4 and Cu(tMe_2D4) bind as intercalators. Primary indicators are the hyperchromic responses and negative CD signals, both diagnostic of intercalative binding (*3*, *4*). Another sign of intercalation is the increase in the specific viscosity of ST DNA that occurs with the uptake of H_2tMe_2D4 or Cu(tMe_2D4) (*52*). The viscosity enhancement occurs because intercalation of a ligand results in an increase in the average length and rigidity of the DNA molecule. In turn, the length increase enhances the resistance to flow as long as the average chain length is compatible with rodlike diffusion (*68*). Last but not least, the adducts formed by Cu(tMe_2D4) and DNA are luminescent. Studies have shown that emission quenching is virtually complete when the analogous Cu(T4) complex binds externally to [poly(dA-dT)]$_2$ (*49*). Intercalation is the only mode of binding that completely blocks access to the axial coordination positions about the copper center. It is interesting in this regard that both the intensity and the wavelength maximum of the emission from the bound form of Cu(tMe_2D4) vary with the DNA host. In particular, the intensity is ∼3 times higher for adduct formation with [poly(dG-dC)]$_2$ than with [poly(dA-dT)]$_2$. A similar emission increase occurs when Cu(T4) migrates to a DNA host with higher melting temperatures (*16*, *62*). A possible explanation is that, during the lifetime of the excited state, structural fluctuations of the host periodically expose the copper center to quenching centers. In that case, the emission intensity varies with the rigidity of the host. When an A═T base pair is part of the intercalation site, free thymine C═O groups are also available to act as nucleophiles.

Intermediate Loading ($2 \leq q \leq 20$). In the intermediate-loading regime when $2 \leq q \leq 20$, the picture is a bit more complicated, but intercalation is still the favored binding motif for H_2tMe_2D4 and Cu(tMe_2D4). The analysis focuses on the results with the copper derivative. To start, the hypochromism induced by the addition of ST DNA, [poly(dG-dC)]$_2$, or [poly(dA-dT)]$_2$ as well as the viscometry data obtained with ST DNA is consistent with intercalation of Cu(tMe_2D4). However, the emission results definitively establish that the binding is by intercalation. The onset of the emission signal coincides with the appearance of [poly(dG-dC)]$_2$ or [poly(dA-dT)]$_2$ in solution, and in both cases, the emission intensity saturates by around $q = 3$. Intercalated Cu(tMe_2D4) is undoubtedly the only contributing species, because neither the emission maximum nor the intensity changes with the addition of excess DNA. If a second emitting species were present, some change in the signal would occur because the emission from Cu(tMe_2D4) is very sensitive to the nature and rigidity of the host, vide supra.

With foreknowledge of the mode of binding, interpretation of the relatively complicated CD becomes feasible. For example, it is easy to understand why the adduct of Cu(tMe_2D4) with [poly(dA-dT)]$_2$ exhibits a multisignate CD spectrum at $q = 3$ (Figure 3), even though the porphyrin binds by intercalation. The problem is that excitonic coupling between neighboring transition moments is unavoidable at such a high loading (*69*). In the literature, there are many reports of bisignate, or bimodal, induced CD signals attributed to self-stacking of cationic porphyrins on the exterior of a polyanionic DNA host (*24*, *29*). However, exciton coupling depends upon proximity, not self-stacking. Elegant studies by Lewis and co-workers (*70*) have demonstrated that stilbene chromophores couple through as many as 10 bp in doubly capped hairpins. Nakanishi and co-workers (*71*) have reported similar effects in solutions containing porphyrin chromophores tethered to opposite ends of DNA duplexes. For a noncovalently bound chromophore, such as Cu(tMe_2D4), the spacing between chromophores may vary, but neighbor−neighbor interactions are inevitable at low q values (high loading) because of the limited pool of available binding sites.

Heavy Loading Regime. Finally, a very different mode of binding occurs at the highest loadings, when $q < 2$. In many respects, this regime is notable for the lack of spectroscopic consequences. Witness, for example, the lack of intensity in the induced CD spectrum and the absence of a significant bathochromic shift in the Soret region. The bound form of Cu(tMe_2D4) is practically nonemissive as well. Indeed, when the DNA base pair-to-porphyrin ratio is small, hyperchromism is the major spectral perturbation, as is the case when porphyrins aggregate in aqueous solution (*29*). When there is too little DNA in solution to take up the ligand, it seems clear that the host simply serves to provide nucleation sites for aggregation of the porphyrin.

Interactions of DNA with Zn(tMe_2D4)

In contrast to Cu(tMe_2D4) and H_2tMe_2D4, Zn(tMe_2D4) apparently exhibits only two phases of binding. Moreover, as the discussion will show, near-neighbor porphyrin−porphyrin interactions persist even in the presence of a large excess of DNA due to cooperative uptake of Zn(tMe_2D4).

Zn(tMe_2D4) and [Poly(dG-dC)]$_2$. Titrating in small amounts of [poly(dG-dC)]$_2$ clearly promotes the aggregation of the Zn(tMe_2D4) porphyrin. Thus, the hypochromism in the Soret band reaches 40% by $q = 10$, despite almost no change in band position and the absence of a significant induced CD signal. By $q = 20$, however, the second, apparently final phase of binding starts to appear. In this loading regime, the hypochromism continues to increase marginally; the new effect is a bathochromic shift of the Soret absorption. When excess [poly(dG-dC)]$_2$ is present, the decrease in hypochromism is very slight, so it seems evident that Zn(tMe_2D4) binds as an intercalator. The LD^r results in Figure 10 support this interpretation. The intercalation of Zn(tMe_2D4) is not so surprising when one recognizes that same experiments reveal that Zn(T4) also intercalates into [poly(dG-dC)]$_2$. The latter observation should draw attention because previous LD studies have found that Zn(T4) binds externally to calf thymus DNA, in a canted orientation relative to the DNA bases (*72*, *73*). One difference is that A=T base pairs outnumber G≡C base pairs in calf thymus DNA, but there is another important qualification to bear in mind as well. Namely, this study used a $\mu = 0.05$ M buffer, and Chirvony et al. (*74*) have already shown that the interaction between Zn(T4) and [poly(dG-dC)]$_2$ is very sensitive to ionic strength. Their analysis of hypochromism and triplet quenching data shows that intercalation is an important binding motif for Zn(T4) at lower ionic strengths ($\mu \approx 0.03-0.05$ M) but that external binding to [poly(dG-dC)]$_2$ becomes dominant at higher ionic strengths.

Two compensating effects account for the fact that uptake by [poly(dG-dC)]$_2$ has little effect on the emission maximum. One is that the zinc porphyrin has to dissociate an axial ligand to bind as an intercalator. If that were the only consideration, the emission would shift to a higher energy (*75*). However, intercalating between bases naturally induces a bathochromic shift of the emission. As demonstrated in previous binding studies involving Zn(D4) (*18*), the net result is little or no shift in the emission maximum. A self-stacked version of four-coordinate, externally bound Zn(tMe_2D4) arguably could give rise to similar absorbance and emission spectra. However, results obtained with the Cu(tMe_2D4) system have established that intercalation is the preferred binding motif for a porphyrin with an in-plane metal center and no axial ligands.

Self-stacking on the [poly(dG-dC)]$_2$ host would also give rise to a CD signal much stronger than the one observed here (*24*). Asymmetric and mainly negative, the induced CD signal of the Zn(tMe_2D4) adduct is clearly bisignate because the CD signal crosses the baseline at essentially the wavelength of the absorption maximum. The CD signal is still compatible with intercalative binding, because the adduct formed by Cu(tMe_2D4) and [poly(dG-dC)]$_2$ exhibits virtually the same band shape, under conditions of intermediate loading. The chromophores have to bind only close enough to each other to experience excitonic coupling. What makes the Zn(tMe_2D4) system unusual is that proximate spacing persists even at very high q values due to a cooperative effect, vide infra.

Zn(tMe_2D4) and [Poly(dA-dT)]$_2$. The zinc porphyrin also exhibits only two phases of binding with [poly(dA-dT)]$_2$. Here, too, the first phase of binding clearly involves aggregation of the porphyrin because the introduction of low

levels of the DNA generates an absorbance decrease but no appreciable shift in the Soret maximum. The approach toward a limiting spectrum begins by $q \approx 10$ base pairs per porphyrin, but Zn(tMe$_2$D4) clearly remains bound externally to [poly(dA-dT)]$_2$. The most telling effect is that the absorption spectrum of the limiting adduct exhibits almost no hypochromism. A four-coordinate form would intercalate into [poly(dA-dT)]$_2$ just as Cu(tMe$_2$D4) does, so Zn(tMe$_2$-D4) must retain a fifth ligand. Nevertheless, some type of structural reorganization occurs, because the emission signal undergoes a distinct hypsochromic shift. The shift in the emission may be an indication of weaker axial ligation (*75*), perhaps because a donor group from the DNA displaces the original water ligand. Kruglik et al. have invoked analogous adduct formation with [poly(dA-dT)]$_2$ to explain Raman results obtained for a photoexcited state of Cu(T4) (*76*). At this juncture, postulating coordinate−covalent bond formation to a DNA base is highly speculative, but that could also explain why irradiation at 260 nm into a short-lived state of [poly(dA-dT)]$_2$ produces emission from Zn(tMe$_2$D4). Previous workers have argued that DNA absorption becomes involved only in the porphyrin excitation spectrum when there is intercalative binding (*77*). In principle, covalent bonding to a DNA base could also give rise to a splitting of the Soret transition and account for the bisignate CD signal (*78*). However, the symmetry of the observed CD spectrum and the fact that the crossover wavelength agrees with the wavelength maximum of the Soret absorption are much more in keeping with a conservative spectrum. As with [poly(dG-dC)]$_2$, the Zn(tMe$_2$D4) ligands must bind close enough to each other on the [poly(dA-dT)]$_2$ host to experience exciton coupling (*78*).

Cooperative Effects. Evidence that points to cooperative binding of Zn(tMe$_2$D4) includes the difference spectrum in Figure 8 as well as with the bisignate CD spectra exhibited by the adducts formed with [poly(dG-dC)]$_2$ and [poly(dA-dT)]$_2$. In the context of cationic porphyrins, the most commonly recognized form of cooperative binding to DNA is self-stacking on the surface of the host (*25*, *29*, *79*). However, more traditional modes of ligand binding can also exhibit cooperativity when binding involves reorganization of the DNA structure (*80*). Whenever extending the reorganization to a neighboring position is easier than initiating a structure change somewhere else in the host, cooperative uptake is a possibility. Relative to Cu(tMe$_2$D4), the reorganization necessary for intercalation of the zinc(II) porphyrin entails an added energy requirement, namely, dissociation of an axial ligand (*18*, *21*). The affinity for Zn(tMe$_2$D4) is therefore weaker, and that may explain why cooperative effects become important. Creation of a cavity requires local unwinding of the host (*81*), and simply extending the structural change to a neighboring site may well facilitate the uptake of another Zn(tMe$_2$D4) ligand. Coulombic repulsion between ligands serves as a countervailing force, but the +2 charge of the ligand should not present much of a problem in the field of the DNA phosphates. Although not much structural information is available, a similar analysis can probably be applied to formation of adducts with [poly(dA-dT)]$_2$ because high-affinity external binding appears to involve substantial structural reorganization (*18*, *19*, *21*). As others have suggested (*82*, *83*), electrostatic effects and/or local hydration may well predispose a sequence that is rich in A=T base pairs to bind bulky porphyrins externally. It is, however, also true that porphyrins are rigid molecules that do not have the crescent shape normally found in ligands tailored for high-affinity groove binding to the canonical DNA structure. Structural reorganization of a flexible host therefore provides an opportunity to maximize interactions with the DNA surface.

CONCLUSIONS AND PERSPECTIVE

Tetrasubstituted H$_2$tMe$_2$D4 behaves more like the disubstituted porphyrin H$_2$D4 than H$_2$T4, because H$_2$tMe$_2$D4 and Cu(tMe$_2$D4) both intercalate into B-form DNA irrespective of its base composition. However, the steric consequences associated with the 5,15-methyl substituents become apparent with the weaker ligand Zn(tMe$_2$D4), which adopts an external binding motif in low-melting, A=T base pair-rich runs of DNA. Compared with H$_2$T4, the dicationic H$_2$tMe$_2$D4 system is more prone to aggregate on the DNA host; however, the problem is minimal by comparison with a really hydrophobic porphyrin like H$_2$T$_{agg}$ (*29*). In order of increasing base pair-to-porphyrin ratios, the three environments that H$_2$tMe$_2$D4 or Cu(tMe$_2$D4) experiences in a typical DNA titration are surface aggregation, followed by densely spaced intercalation, and then dispersion into independent, well-separated intercalation sites. Uptake of Zn(tMe$_2$D4) is very different because binding is a strictly cooperative process and the motif depends on the base content of the DNA host. Heretofore, external stacking has been widely recognized as a binding motif compatible with cooperative uptake of a cationic porphyrin (*22*, *56*, *79*), but cooperative binding may occur whenever substantial structural reorganization is necessary. Recognizing when cooperative binding occurs is important because it impacts everything from competitive binding assays (*84*) to comparative kinetics studies (*85*, *86*).

In the case of porphyrins, through-space or through-bond interactions involving near neighbors complicate spectral analyses when the chromophores do not resonate independently. For example, one cannot assume that an intercalated porphyrin will give a negative induced CD signal in the Soret region, when other chromophores bind nearby. The reason is that coupling between chromophores produces a multisignate response that depends on the distance(s) of separation and therefore the loading. While dipole−dipole coupling is an excited state interaction, alteration of the spectral properties can complicate the process of extracting ground state information, such as the binding constant. Fortunately, dipole−dipole interactions do not affect the luminescence studies of copper(II) porphyrins due to the forbidden character of the transition. In DNA binding studies with Cu(tMe$_2$D4), data from emission and CD studies yield unusually detailed information about the intermediate loading regime. The emission results are notable because they establish that Cu(tMe$_2$D4) binds by intercalation and that the porphyrin has a footprint of ~3 bp. Juxtapositional metal-centered and $\pi-\pi^*$ excited states render the emission of a copper(II) porphyrin uniquely sensitive to the presence of Lewis bases (*4*, *21*, *49*, *62*). To cite examples, Lewis bases do not quench the emission of a palladium(II) porphyrin because the metal-centered states are inaccessible, while nickel(II) porphyrins are nonemissive because metal-centered excited states occur at relatively low energies (*87*, *88*).

REFERENCES

1. Fiel, R. J. (1989) Porphyrin-Nucleic Acid Interactions: A Review, *J. Biomol. Struct. Dyn. 6* (6), 1259−1275.
2. Marzilli, L. G. (1990) Medical Aspects of DNA-Porphyrin Interactions, *New J. Chem. 14* (6−7), 409−420.
3. Pasternack, R. F., and Gibbs, E. J. (1996) Porphyrin and metalloporphyrin interactions with nucleic acids, *Met. Ions Biol. Syst. 33*, 367−397.
4. McMillin, D. R., and McNett, K. M. (1998) Photoprocesses of copper complexes that bind to DNA, *Chem. Rev. 98* (3), 1201−1219.
5. Kessel, D., Luguya, R., and Vicente, M. G. H. (2003) Localization and photodynamic efficacy of two cationic porphyrins varying in charge distribution, *Photochem. Photobiol. 78* (5), 431−435.
6. Lang, K., Mosinger, J., and Wagnerova, D. M. (2004) Photophysical properties of porphyrinoid sensitizers non-covalently bound to host molecules: Models for photodynamic therapy, *Coord. Chem. Rev. 248* (3−4), 321−350.
7. Griffiths, J. (2004) Colourful Therapy, *Educ. Chem. 41* (3), 71−73.
8. Han, F. X. G., Wheelhouse, R. T., and Hurley, L. H. (1999) Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition, *J. Am. Chem. Soc. 121* (15), 3561−3570.
9. Cech, T. R. (2000) Life at the end of the chromosome: Telomeres and telomerase, *Angew. Chem., Int. Ed. 39* (1), 35−43.
10. Carvlin, M. J., Dattagupta, N., and Fiel, R. J. (1982) Circular-Dichroism Spectroscopy of a Cationic Porphyrin Bound to DNA, *Biochem. Biophys. Res. Commun. 108* (1), 66−73.
11. Fiel, R. J., Howard, J. C., Mark, E. H., and Dattagupta, N. (1979) Interaction of DNA with a Porphyrin Ligand: Evidence for Intercalation, *Nucleic Acids Res. 6* (9), 3093−3118.
12. Carvlin, M. J., and Fiel, R. J. (1983) Intercalative and Nonintercalative Binding of Large Cationic Porphyrin Ligands to Calf Thymus DNA, *Nucleic Acids Res. 11* (17), 6121−6139.
13. Fiel, R. J., and Munson, B. R. (1980) Binding of Meso-Tetra(4-N-Methylpyridyl) Porphine to DNA, *Nucleic Acids Res. 8* (12), 2835−2842.
14. Pasternack, R. F., Gibbs, E. J., and Villafranca, J. J. (1983) Interactions of Porphyrins with Nucleic-Acids, *Biochemistry 22* (23), 5409−5417.
15. Strickland, J. A., Marzilli, L. G., and Wilson, W. D. (1990) Binding of Meso-Tetrakis(N-Methylpyridiniumyl)Porphyrin Isomers to DNA: Quantitative Comparison of the Influence of Charge-Distribution and Copper(II) Derivatization, *Biopolymers 29* (8−9), 1307−1323.
16. Eggleston, M. K., Crites, D. K., and McMillin, D. R. (1998) Studies of the base-dependent binding of Cu(T4) to DNA hairpins (H_2T4 = meso-tetrakis(4-(N-methylpyridiumyl))porphyrin), *J. Phys. Chem. A 102* (28), 5506−5511.
17. Gibbs, E. J., Maurer, M. C., Zhang, J. H., Reiff, W. M., Hill, D. T., Malickablaszkiewicz, M., McKinnie, R. E., Liu, H. Q., and Pasternack, R. F. (1988) Interactions of Porphyrins with Purified DNA and More Highly Organized Structures, *J. Inorg. Biochem. 32* (1), 39−65.
18. Bejune, S. A., Shelton, A. H., and McMillin, D. R. (2003) New Dicationic Porphyrin Ligands Suited for Intercalation into B-Form DNA, *Inorg. Chem. 42*, 8465−8475.
19. Tears, D. K. C., and McMillin, D. R. (1998) Duplex hydrogen banding promotes intercalation of Cu(T4) in DNA hairpins (Cu(T4) = meso-tetrakis(4-(N-methylpyridyl))porphyrincopper(II)), *Chem. Commun. 22*, 2517−2518.
20. Strickland, J. A., Marzilli, L. G., Gay, K. M., and Wilson, W. D. (1988) Porphyrin and Metalloporphyrin Binding to DNA Polymers: Rate and Equilibrium Binding Studies, *Biochemistry 27* (24), 8870−8878.
21. McMillin, D. R., Shelton, A. H., Bejune, S. A., Fanwick, P. E., and Wall, R. K. (2005) Understanding binding interactions of cationic porphyrins with B-form DNA, *Coord. Chem. Rev. 249* (13−14), 1451−1459.
22. Gibbs, E. J., Tinoco, I., Maestre, M. F., Ellinas, P. A., and Pasternack, R. F. (1988) Self-Assembly of Porphyrins on Nucleic-Acid Templates, *Biochem. Biophys. Res. Commun. 157* (1), 350−358.
23. Sari, M. A., Battioni, J. P., Dupre, D., Mansuy, D., and Lepecq, J. B. (1990) Interaction of Cationic Porphyrins with DNA: Importance of the Number and Position of the Charges and Minimum Structural Requirements for Intercalation, *Biochemistry 29* (17), 4205−4215.
24. Pasternack, R. F., Ewen, S., Rao, A., Meyer, A. S., Freedman, M. A., Collings, P. J., Frey, S. L., Ranen, M. C., and de Paula, J. C. (2001) Interactions of copper(Il) porphyrins with DNA, *Inorg. Chim. Acta 317* (1−2), 59−71.
25. Pasternack, R. F., Gibbs, E. J., and Villafranca, J. J. (1983) Interactions of Porphyrins with Nucleic Acids, *Biochemistry 22* (10), 2406−2414.
26. Lipscomb, L. A., Zhou, F. X., Presnell, S. R., Woo, R. J., Peek, M. E., Plaskon, R. R., and Williams, L. D. (1996) Structure of a DNA-porphyrin complex, *Biochemistry 35* (9), 2818−2823.
27. Guliaev, A. B., and Leontis, N. B. (1999) Cationic 5,10,15,20-tetrakis(N-methylpyridinium-4-yl)porphyrin fully intercalates at 5′-CG-3′ steps of duplex DNA in solution, *Biochemistry 38* (47), 15425−15437.
28. Wall, R. K., Shelton, A. H., Bonaccorsi, L. C., Bejune, S. A., Dube, D., and McMillin, D. R. (2001) H_2D3: A cationic porphyrin designed to intercalate into B-form DNA (H_2D3 = trans-di(N-methylpyridium-3-yl)porphyrin), *J. Am. Chem. Soc. 123* (46), 11480−11481.
29. Pasternack, R. F., Bustamante, C., Collings, P. J., Giannetto, A., and Gibbs, E. J. (1993) Porphyrin Assemblies on DNA as Studied by a Resonance Light-Scattering Technique, *J. Am. Chem. Soc. 115* (13), 5393−5399.
30. Lin, V. S. Y., Iovine, P. M., DiMagno, S. G., and Therien, M. J. (2002) Dipyrryl and porphyrinic precursors to supramolecular conjugated (porphinato)metal arrays: Synthesis of dipyrrylmethane and (5,15-diphenylporphinato)zinc(II), *Inorg. Synth. 33*, 55−61.
31. Manka, J. S., and Lawrence, D. S. (1989) High-Yield Synthesis of 5,15-Diarylporphyrins, *Tetrahedron Lett. 30* (50), 6989−6992.
32. Wang, Q. M., and Bruce, D. W. (1995) One-Step Synthesis of Beta,Meso-Unsubstituted Dipyrromethane, *Synlett 12*, 1267−1268.
33. Batinic-Haberle, I., Spasojevic, I., Hambright, P., Benov, L., Crumbliss, A. L., and Fridovich, I. (1999) Relationship among redox potentials, proton dissociation constants of pyrrolic nitrogens, and in vivo and in vitro superoxide dismutating activities of manganese(III) and iron(III) water-soluble porphyrins, *Inorg. Chem. 38* (18), 4011−4022.
34. Sambrook, J., Fritsch, J., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Springs Harbor Laboratory Press, Plainview, NY.
35. Inman, R. B., and Baldwin, R. L. (1962) Helix-Random Coil Transitions in Synthetic DNAs of Alternating Sequence, *J. Mol. Biol. 5*, 172−184.
36. Grant, R. C., Wells, R. D., and Kodama, M. (1972) Enzymatic and Physical Studies on $(dI\text{-}dC)_n\cdot(dI\text{-}dC)_n$ and $(dG\text{-}dC)_n\cdot(dG\text{-}dC)_n$, *Biochemistry 11* (5), 805−815.
37. Felsenfeld, G., and Hirschman, S. Z. (1965) A Neighbor-Interaction Analysis of the Hypochromism and Spectra of DNA, *J. Mol. Biol. 13*, 407−427.
38. Satyanarayana, S., Dabrowiak, J. C., and Chaires, J. B. (1992) Neither Δ-Tris(phenanthroline)ruthenium(II) nor Λ-Tris(phenanthroline)ruthenium(II) Binds to DNA by Classical Intercalation, *Biochemistry 31* (39), 9319−9324.
39. Rodger, A. (1993) Linear Dichroism, *Methods Enzymol. 226*, 232−258.
40. Rodger, A., and Norden, B. (1997) *Circular Dichroism and Linear Dichroism*, Oxford University Press, New York.
41. Marrington, R., Dafforn, T. R., Halsall, D. J., MacDonald, J. I., Hicks, M., and Rodger, A. (2005) Validation of new microvolume Couette flow linear dichroism cells, *Analyst 130*, 1608−1616.
42. Lindsey, J. S. (1999) Synthesis of meso-Substituted Porphyrins, in *The Porphyrin Handbook* (Kadish, K. M., Smith, K. M., and Guilard, R., Eds.) pp 45−118, Academic Press, New York.
43. Vicente, M. G. H., and Smith, K. M. (2000) Porphyrins and derivatives: Synthetic strategies and reactivity profiles, *Curr. Org. Chem. 4*, 139−174.
44. Longo, F. R., Thorne, E. J., Adler, A. D., and Dym, S. (1975) Notes on the Synthesis of Porphin, *J. Heterocycl. Chem. 12*, 1305−1309.
45. Kim, J. B., Adler, A. D., and Longo, F. R. (1978) Synthesis of Porphyrins from Monopyrroles, in *The Porphyrins* (Dolphin, D., Ed.) pp 85−100, Academic Press, New York.
46. Goncalves, D. P. N., Ladame, S., Balasubramanian, S., and Sanders, J. K. M. (2006) Synthesis and G-quadruplex binding studies of new 4-N-methylpyridinium porphyrins, *Org. Biomol. Chem. 4*, 3337−3342.

47. Asano-Someda, M., and Kaizu, Y. (1995) Red-Shifted Emission Spectra of Several Mesosubstituted Copper Porphyrins in Fluid Solution, *J. Photochem. Photobiol., A 87*, 23−29.
48. Shah, B., and Hambright, P. (1970) Acid Catalyzed Solvolysis Reactions of Zinc Porphyrins, *J. Inorg. Nucl. Chem. 32*, 3420−3422.
49. Hudson, B. P., Sou, J., Berger, D. J., and McMillin, D. R. (1992) Luminescence Studies of the Intercalation of Cu(TMPyP4) into DNA, *J. Am. Chem. Soc. 114*, 8997−9002.
50. Mojzes, P., Chinsky, L., and Turpin, P. Y. (1993) Interaction of Electronically Excited Copper(II) Porphyrin with Oligonucleotides and Polynucleotides: Exciplex Building Process by Photoinitiated Axial Ligation of Porphyrin to Thymine and Uracil Residues, *J. Phys. Chem. 97*, 4841−4847.
51. Saucier, J. M., Lepecq, J. B., and Festy, B. (1971) Change of Torsion of DNA Helix Caused by Intercalation. 2. Measurement of Relative Change of Torsion Induced by Various Intercalating Drugs, *Biochimie 53*, 973−980.
52. Banville, D. L., Marzilli, L. G., Strickland, J. A., and Wilson, W. D. (1986) Comparison of the Effects of Cationic Porphyrins on DNA Properties: Influence of GC Content of Native and Synthetic Polymers, *Biopolymers 25*, 1837−1858.
53. Tjahjono, D. H., Kartasasmita, R. E., Nawawi, A., Mima, S., Akutsu, T., Yoshioka, N., and Inoue, H. (2006) Binding of tetrakis-(pyrazoliumyl)porphyrin and its copper(II) and zinc(II) complexes to poly(dG-dC)$_2$ and poly(dA-dT)$_2$, *J. Biol. Inorg. Chem. 11*, 527−538.
54. Pasternack, R. F., Centuro, G. C., Boyd, P., Hinds, L. D., Huber, P. R., Francesc, L., Fasella, P., Engasser, G., and Gibbs, E. (1972) Aggregation of Meso-Substituted Water-Soluble Porphyrins, *J. Am. Chem. Soc. 94*, 4511−4517.
55. Kano, K., Fukuda, K., Wakami, H., Nishiyabu, R., and Pasternack, R. F. (2000) Factors influencing self-aggregation tendencies of cationic porphyrins in aqueous solution, *J. Am. Chem. Soc. 122*, 7494−7502.
56. Pasternack, R. F., Giannetto, A., Pagano, P., and Gibbs, E. J. (1991) Self-Assembly of Porphyrins on Nucleic Acids and Polypeptides, *J. Am. Chem. Soc. 113*, 7799−7800.
57. Szintay, G., Horvath, A., and Grampp, G. (1999) Temperature dependence study of pyridine complex formation and emission quenching of copper(II) octaethyl- and tetraphenylporphyrin, *J. Photochem. Photobiol., A 126*, 83−89.
58. Gouterman, M. (1978) Optical Spectra and Electronic Structure of Porphyrins and Related Rings, in *The Porphyrins* (Dolphin, D., Ed.) pp 1−165, Academic Press, New York.
59. Kim, D., Holten, D., and Gouterman, M. (1984) Evidence from Picosecond Transient Absorption and Kinetic Studies of Charge-Transfer States in Copper(II) Porphyrins, *J. Am. Chem. Soc. 106*, 2793−2798.
60. Cunningham, K. L., McNett, K. M., Pierce, R. A., Davis, K. A., Harris, H. H., Falck, D. M., and McMillin, D. R. (1997) EPR spectra, luminescence data, and radiationless decay processes of copper(II) porphyrins, *Inorg. Chem. 36*, 608−613.
61. Liu, F., Cunningham, K. L., Uphues, W., Fink, G. W., Schmolt, J., and McMillin, D. R. (1995) Luminescence Quenching of Copper(II) Porphyrins with Lewis Bases, *Inorg. Chem. 34*, 2015−2018.
62. Lugo-Ponce, P., and McMillin, D. R. (2000) DNA-binding studies of Cu(T4), a bulky cationic porphyrin, *Coord. Chem. Rev. 208*, 169−191.
63. Chen, L. X. (2004) Taking snapshots of photoexcited molecules in disordered media by using pulsed synchrotron X-rays, *Angew. Chem., Int. Ed. 43*, 2886−2905.
64. Kruglik, S. G., Apanasevich, P. A., Chirvony, V. S., Kvach, V. V., and Orlovich, V. A. (1995) Resonance Raman, Cars, and Picosecond Absorption Spectroscopy of Copper Porphyrins: The Evidence for the Exciplex Formation with Oxygen-Containing Solvent Molecules, *J. Phys. Chem. 99*, 2978−2995.
65. Chirvony, V. S. (2003) Primary photoprocesses in cationic 5,10,-15,20-meso-tetrakis(4-N-methylpyridiniumyl)porphyrin and its transition metal complexes bound with nucleic acids, *J. Porphyrins Phthalocyanines 7*, 766−774.
66. Kruglik, S. G., Mojzes, P., Mizutani, Y., Kitagawa, T., and Turpin, P. Y. (2001) Time-resolved resonance raman study of the exciplex formed between excited Cu-porphyrin and DNA, *J. Phys. Chem. B 105*, 5018−5031.
67. Keating, L. R., and Szalai, V. A. (2004) Parallel-stranded guanine quadruplex interactions with a copper cationic porphyrin, *Biochemistry 43*, 15891−15900.
68. Muller, W., and Crothers, D. M. (1968) Studies of Binding of Actinomycin and Related Compounds to DNA, *J. Mol. Biol. 35*, 251−290.
69. Huang, X. F., Nakanishi, K., and Berova, N. (2000) Porphyrins and metalloporphyrins: Versatile circular dichroic reporter groups for structural studies, *Chirality 12*, 237−255.
70. Lewis, F. D. (2005) DNA molecular photonics, *Photochem. Photobiol. 81*, 65−72.
71. Balaz, M., Holmes, A. E., Benedetti, M., Rodriguez, P. C., Berova, N., Nakanishi, K., and Proni, G. (2005) Synthesis and circular dichroism of tetraarylporphyrin-oligonucleotide conjugates, *J. Am. Chem. Soc. 127*, 4172−4173.
72. Geacintov, N. E., Ibanez, V., Rougee, M., and Bensasson, R. V. (1987) Orientation and Linear Dichroism Characteristics of Porphyrin DNA Complexes, *Biochemistry 26*, 3087−3092.
73. Strickland, J. A., Banville, D. L., Wilson, W. D., and Marzilli, L. G. (1987) Metalloporphyrin Effects on Properties of DNA Polymers, *Inorg. Chem. 26*, 3398−3406.
74. Chirvony, V. S., Galievsky, V. A., Terekhov, S. N., Dzhagarov, B. M., Ermolenkov, V. V., and Turpin, P. Y. (1999) Binding of the Cationic 5-Coordinate Zn(II)-5,10,15,20-tetrakis(4-*N*-methylpyridyl)porphyrin to DNA and Model Polynucleotides: Ionic-Strength Dependent Intercalation in [Poly(dG-dC)]$_2$, *Biospectroscopy 5*, 302−312.
75. Whitten, D. G., Lopp, I. G., and Wildes, P. D. (1968) Fluorescence of Zinc and Magnesium Etioporphyrin. I. Quenching and Wavelength Shifts Due to Complex Formation, *J. Am. Chem. Soc. 90*, 7196−7200.
76. Kruglik, S. G., Galievsky, V. A., Chirvony, V. S., Apanasevich, P. A., Ermolenkov, V. V., Orlovich, V. A., Chinsky, L., and Turpin, P. Y. (1995) Dynamics and Mechanism of the Exciplex Formation between Cu(Tmpy-P4) and DNA Model Compounds Revealed by Time-Resolved Transient Absorption and Resonance Raman Spectroscopies, *J. Phys. Chem. 99*, 5732−5741.
77. Sari, M. A., Battioni, J. P., Mansuy, D., and Lepecq, J. B. (1986) Mode of Interaction and Apparent Binding Constants of Meso-Tetraaryl Porphyrins Bearing between One and Four Positive Charges with DNA, *Biochem. Biophys. Res. Commun. 141*, 643−649.
78. Berova, N., and Nakanishi, K. (2000) Exciton Chirality Method: Principles and Applications, in *Circular Dichroism: Principles and Applications* (Nakanishi, K., Berova, N., and Woody, R., Eds.) pp 337−382, Wiley-VCH, New York.
79. Armitage, B. A. (2006) DNA-Templated Assembly of Helical Multichromophore Aggregates, *Mol. Supramol. Photochem. 14*, 255−287.
80. Bloomfield, W. A., Crothers, D. M., and Tinoco, J., Jr. (2000) *Nucleic Acids: Structures, Properties, and Functions*, University Science, Sausalito, CA.
81. Calladine, C. R., and Drew, H. R. (1997) *Understanding DNA*, 2nd ed., Academic Press, New York.
82. Marzilli, L. G., Petho, G., Lin, M. F., Kim, M. S., and Dixon, D. W. (1992) Tentacle Porphyrins-DNA Interactions, *J. Am. Chem. Soc. 114*, 7575−7577.
83. Mukundan, N. E., Petho, G., Dixon, D. W., and Marzilli, L. G. (1995) DNA-Tentacle Porphyrin Interactions: AT Over GC Selectivity Exhibited by an Outside Binding Self-Stacking Porphyrin, *Inorg. Chem. 34*, 3677−3687.
84. Pasternack, R. F., Caccam, M., Keogh, B., Stephenson, T. A., Williams, A. P., and Gibbs, E. J. (1991) Long-Range Fluorescence Quenching of Ethidium Ion by Cationic Porphyrins in the Presence of DNA, *J. Am. Chem. Soc. 113*, 6835−6840.
85. Lincoln, P., Tuite, E., and Norden, B. (1997) Short-circuiting the molecular wire: Cooperative binding of Δ-[Ru(phen)$_2$dppz]$^{2+}$ and Δ-[Rh(phi)$_2$bipy]$^{3+}$ to DNA, *J. Am. Chem. Soc. 119*, 1454−1455.
86. Olson, E. J. C., Hu, D. H., Hormann, A., and Barbara, P. F. (1997) Quantitative modeling of DNA-mediated electron transfer between metallointercalators, *J. Phys. Chem. B 101*, 299−303.
87. Rodriguez, J., and Holten, D. (1990) Ultrafast Photodissociation of a Metalloporphyrin in the Condensed Phase, *J. Chem. Phys. 92*, 5944−5950.
88. Kruglik, S. G., Mizutani, Y., and Kitagawa, T. (1997) Time-resolved resonance Raman study of the primary photoprocesses of nickel(II) octaethylporphyrin in solution, *Chem. Phys. Lett. 266*, 283−289.