MOAC PhD Proposals

2007 Academic Year

MOAC PhD-project proposal

Project title: Bacterial cell division: molecular level understanding of the early stages of the division ring formation

Supervisor name	Department/email/phone number	Supervisor's advisor
Alison Rodger (AR)	Chemistry/a.rodger@warwick.ac.uk/74696/07876218199	N/A
David Roper (DR)	Biological Sciences/david.roper@warwick.ac.uk/28369	N/A
Matthew Turner (MT)	Physics/m.s.turner@warwick.ac.uk/22257	N/A

and action of the Z-ring.

Project proposal:

1. **Track record of supervisors in research and supervision**

AR: has supervised 31 PhD students since 1991 of which 22 have successfully completed their PhD. She has 7 students currently, all of whom are co-supervised by internal or external supervisors. Two of the 31 students withdrew after 1 year. All PhD students have published at least 2 papers on their work. AR is internationally recognized as an expert in the development and application of polarized UV/visible light spectroscopies to study the structure and function of biomacromolecules. She has published 6 books, 9 book chapters and over 125 refereed papers. She currently has one EPSRC research project grant, a BBSRC Research Equipment Initiative grant, is the PI on the MOAC grants and a co-I on the Systems Biology DTC. The DTC programmes undertake graduate training in an innovative way; AR has this project.

DR: Has supervised 7 PhD students since 2001 three of whom have now successfully completed their PhD and in addition has supervised 1 overseas MSc student who has successfully completed the thesis. He has 4 PhD students currently registered and working in his laboratory. DR is internationally recognized as a structural biologist in the area of antibiotic resistance and cell division and has published 47 refereed papers and contributed to 1 book chapter. DR currently has an MRC collaboration grant award along with co-applicant status on, Wellcome Trust equipment, BBSRC Research Equipment Initiative grants along with MRC and EPSRC project grants.

MT: Is a theoretical scientist with a background in soft condensed matter and biological physics. His recent interests include the self-assembly and structure of biological fibers, molecular motors, membrane dynamics and structure and genetic networks. He has supervised 6 PhD students, of which two are current and one withdrew on ground of ill health. His students generate on average 3 or 4 papers during their PhD and his last PhD student to graduate won *Young geneticist of the year* and secured a prestigious EMBO fellowship. He has published more than 60 papers generating 1000+ citations with an h-index of 18.

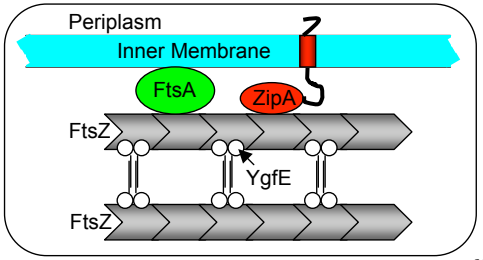
2. **Background to research project**

AIM: The overall aim of this project is to use *in vitro* reconstitution combined with a novel biophysical technique, linear dichroism, to provide a molecular level understanding of the early events of bacterial cell division. To achieve this aim we have subdivided the project into 3 work packets which concern FtsZ in the presence of: 1) the FtsZ bundling protein, YgfE; 2) the integral membrane FtsZ anchor, ZipA; and 3) the peripheral membrane FtsZ anchor, FtsA.

The cell division process in *E. coli* has the potential to provide information of wide applicability about the mechanisms whereby cytoskeletal elements coordinate to alter cellular morphology. Cell division involves interactions between proteins in fibres and proteins in membranes which among other things lead to changes in cell membrane morphology (ultimately one membrane becoming two). These interactions are comparatively poorly understood, and until we do, we cannot control or exploit them *e.g.* to develop antibiotic drugs. In this project *E. coli* cell division will be used as a model process. We already know with some certainty from genetic studies the complete protein interaction pathway for division. In addition the production of GFP fusions of many of the key proteins has allowed their spatial distribution to be tracked through the division process. The structures of many of the soluble proteins involved in the division process have also been solved and these have added significantly to our understanding of some of the protein-protein interactions during division. The stage is thus set for a molecular level understanding of the concerted division process to be achieved. The aim of this multidisciplinary project is to established reliable protein production methods for *E. coli* cell division proteins and their mutants and then develop and apply biophysical techniques to study molecular aspects of cell division that have not been studied easily by pre-existing techniques. The aim is to determine how polymeric proteins of the cytoskeleton interact together and how these complexes may alter membranes morphology.

***E. coli* cell division**

The remodelling of the *E. coli* membrane and peptidoglycan wall during cell division entails the formation of a biomolecular complex involving upwards of 10 proteins which coordinate the curvature of the membrane to form the septum which becomes the ends of the two daughter cells. In the initial stages of bacterial cell division a ring of protein (FtsZ) forms around the narrow axis in the centre of bacterial cell under the inner cell membrane. Other proteins are then recruited to this ring to form a macro structure known as the divisome that, it is thought, may mediate membrane morphology changes during division. For a review see [23]. This process is fundamental for bacterial division, however, as yet we do not understand it well enough to control it.



The early stages of division ring formation

The interactions at the division ring have to be choreographed in order to lead to cell successful division. Recent work *in vivo* has led to the hypothesis that the control of these complex associations is not wholly centred around control of expression (as originally supposed [24]) but may, instead, be determined by alosteric changes within sub-complexes. In this project we will study the first events of cell division which involve the formation of complexes of 4 proteins (FtsZ, FtsA, ZipA and ZapA) on the inner surface of the inner cell membrane. It is known that FtsZ, a structural homologue of tubulin, is the major polymerising element in the complex; it first binds GTP then oligomerises to form a filament which encircles the bacterium forming the Z ring. We have followed this process using LD and thus acquired new structural and mechanistic information about the process [18, 19, 21]**.** To date our LD data have been qualitatively interpreted, we now propose to improve this to provide quantitative complements to data from a range of other biophysics techniques as discussed below. In *E. coli* the co-protein ZapA (also known as YgfE) binds to the Z ring and has been shown *in vitro* to stabilise and thicken the fibrous form of FtsZ. Two further co-proteins, ZipA and FtsA, mediate (in some fashion) the interaction between the FtsZ ring and the inner membrane in *E. coli.* ZipA is thought to be tethered to the membrane by a single N-terminal trans-membrane region which is linked via flexible region to a C-terminal FtsZ binding domain. ZipA binds to FtsZ filaments inducing them to bundle and tethering them to the membrane. FtsA is more loosely associated with the membrane as a dimer via a C-terminal amphipathic helix. Thus far, no FtsZ bundling activity has been attributed to FtsA, although it is known to bind to ATP bind and hydrolyse and has been implicated as a "motor protein" [25] in bacterial division.

3. **Research project**

WP1. How does YgfE bundle FtsZ?

This phase of the project will provide a definitive picture of how YgfE bundles FtsZ filaments. In particular determining whether bundling is induced by one or more of the following processes: 1) an inhibition of GTP hydrolysis; 2) a conformational/orientation change in the fibre; and/or GTP 3) a direct molecular linkage between fibres mediated by YgfE.

WP2. How does the integral membrane anchor ZipA alter FtsZ filament structure and dynamics?

Our previous work on YgfE and its bundling of FtsZ has shown that we can provide a picture of the FtsZ fibre bundling. WP1 will produce a molecular level model of YgfE-FtsZ. Using a similar approach with ZipA we will produce a model that can be compared and contrasted with that for YgfE. These data will help us understand the different roles of YgfE and ZipA in divisome formation. The presence of membrane will offer a further exciting element that will allow us, for the first time, assess the influence of membrane tethering on FtsZ filament formation and membrane morphology

WP3. How does the surface associated membrane anchor FtsA alter FtsZ filament structure and dynamics?

This study of FtsA will provide details of the interaction between FtsA and FtsZ at a molecular level when FtsA is both free in solution and membrane bound. Comparison of these details with those for YgfE and ZipA will, for the first time, allow us to propose specific functions for each protein in divisome formation.

4. **Scientific strategy**

The scientific strategy of this project involve building a molecular model of the *E. Coli* divisome to understand the cytoskeletal organization that leads to cell division. This is part of the wider challenge of understanding how bacterial cells organize their molecular components. Understanding this is the key to the development of antibacterial agents.

5. **Technical training**

The project will involve the student mastering molecular biology techniques, a wide range of biophysics (including linear dichroism, circular dichroism, fluorescence, absorption, kinetics, light scattering, isothermal calorimetry, analytical ultracentrifugation, gel techniques), as well as mathematical and computational modeling. The supervisors will either rprovide or arrange the required training.

6. **Integration of the multi-disciplinary content**

AR and DR have collaborated on FtsZ projects for a number of years, to date this has led to 5 publications. They have co-supervised a successful PhD project. MT has worked on a range of fibrous proteins and has played a key role on advisory committees for AR/DR's students. The team will have monthly meetings to discuss the project and the student will meet with individual supervisors more frequently.

7. Justification of resources required

The project will involve significant molecular biology to produce the required proteins. Therefore the maximum allowable budget will not cover the costs.

8. Project management

The project will be managed by the supervisors via the monthly meetings with the student delegated the responsibility of organizing these meetings. The student will produce a 1 page report for each such meeting as well as any data they wish to present. The student will meet with one of their supervisors every 1–2 days depending on what phase of the project they are currently engaged in.

9. Plan for what the student will do in the first 3−6 months

In a previous study we have uncovered much of the mechanism of YgfE induce FtsZ bundling, however, how these events lead to filament association and which part of YgfE is responsible for bundling is still very much a mystery. Our AUC studies have shown that YgfE exists in a dimer:tetramer equilibrium. This is consistent with a crystal structures of a homologue of YgfE (see figure 2) which shows that a dimer is formed by the association of a two α-helices (one from each monomer) to form a coiled-coil; an end to end association of the coiled coil domains could lead to a "dog bone" shaped tetramer with globular domains at each end. This has led to the hypothesis that YgfE could induce bundling by linking FtsZ filaments like rungs in a filament ladder. In WP1 we will to test this hypothesis using YgfE truncates and a range of techniques. Three truncated YgfE molecules will be constructed:

Figure 2 the crystal structure of a dimer of ZapA (a structural homologue of YgfE) [1]

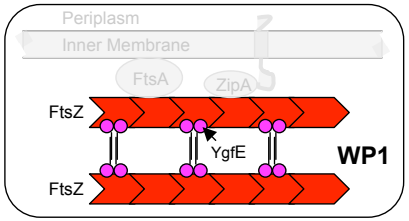
(i) A constitutive YgfE dimer containing the N-terminal globular domain and 8 turns of the coiled coil domain, hence removing the contacts that stabilise the tetramer.

(ii)A constitutive YgfE Monomer in which the coiled-coil domain will be deleted thus removing a significant part of the monomer:monomer interface. This will allow us to assess two aspects of YgfE. Firstly, is the coiled coil domain a requirement for dimerisation of YgfE? If it is then we will also be able to test whether the globular N-terminus can induce FtsZ bundling. If YgfE is still dimeric we will be able to test whether the coiled coil is important for YgfE activity.

(iii)The coiled-coil domain of YgfE will be used to test whether this domain is important in FtsZ interactions and bundling.

Producing and testing YgfE constructs

Each YgfE construct is <40 amino acids and will be produced by peptide synthesis, thus eliminating any issues with production of such small, potentially active proteins in *E. coli*. The oligomerisation state and fold of each construct will then be tested using analytical ultracentrifugation (AUC) and circular dichroism (CD) respectively. The effect of each construct on FtsZ filaments will then be tested using techniques already established in our laboratory to answer the following questions:



A summary of the molecular interactions studier in this WP

WP1.1) Does the construct induce FtsZ bundling? This will initially be tested by measuring changes in the light scattering signal that is produced when FtsZ polymerises. An increases in scattering indicates FtsZ filament bundling. However light scattering is insensitive to the morphology of the filament. We will therefore use LD to look at a number of fundamental aspects of the structure of the filament. LD will probe the order of the filaments, as disruption of fibre order leads to a decrease in LD signal while ordered bundling leads to an enhancement. LD will also provide details of any rearrangements of the subunits in the filament. These changes are manifest in a change in LD spectral shape. In addition we will follow any changes in the conformation of the GTP bound to the filament by monitoring the near-UV guanine LD signal. The LD spectra produced in this study will be deconvolved (using software that we have recently developed as part of the EPSRC project (D075750) with Prof. J. Hirst, Nottingham, http://comp.chem.nottingham.ac.uk/dichrocalc/) to produce direct measures of geometry changes in the fibre. Data of this type will be used in combination with the known X-ray structures of the monomeric FtsZ and the YgfE homologue to construct models of the fibre which will, for the first time, provide a molecular description of how fibre bundling is induced.

WP1.2) Does the construct alter GTP hydrolysis? Hydrolysis of GTP by the FtsZ filament is an inherent part of its dynamic nature. Using a continuous phosphate detection assay we have already developed, we will probe the effect of the constructs on the GTP hydrolysis rate to understand what part of YgfE inhibits hydrolysis.

WP1.3) Does the construct alter the macro structure of the filaments? YgfE induced bundling of FtsZ is easily visible using negative stain TEM. FtsZ fibres formed in the presence of YgfE truncates will be examined and compared with LD data. In addition to this analysis of structure the student will undertake further crystallisation and heavy metal soaking experiments to complete an X-ray structure determination of E.coli YgfE. Attempts to solve the structure with our existing high resolution X-ray data for YgfE using the ZapA as a search model have been unsuccessful providing strong evidence that there are significant structural differences between YgfE and ZapA. These differences may have important functional significance that is crucial to understanding the role of YgfE. Given our ability to produce and control the FtsZ polymerisation a future structural target for this student would be a complex of YgfE and FtsZ.

Project no. 2

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Monday 16th June, 2008)

Project title: Quantitative Enzymology of the Bacterial Penicillin Binding Proteins

Supervisor name	Department/email/phone number	Supervisor's advisor*
Prof Timothy DH Bugg	Department of Chemistry Email T.D.Bugg@warwick.ac.uk, Tel x 73018	
Prof Christopher G Dowson	Department of Biological Sciences Email C.G.Dowson@warwick.ac.uk, Tel x 23534	
Dr David I Roper	Department of Biological Sciences Email David.Roper@warwick.ac.uk, Tel x 28369	

Project proposal:

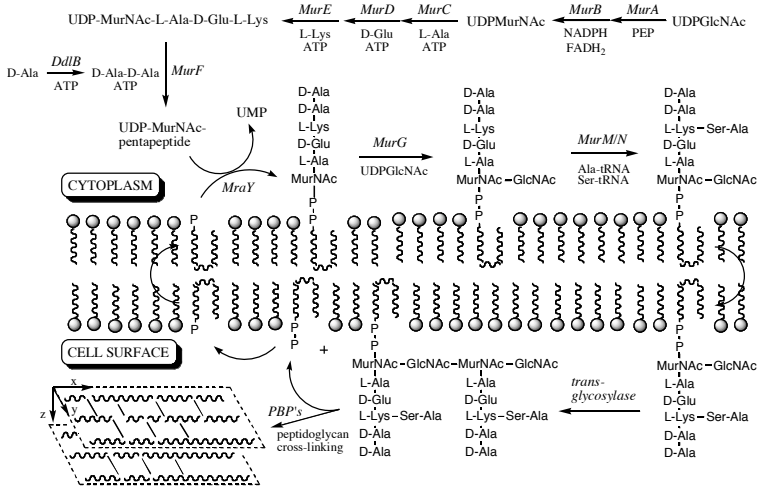
1. Track record of supervisors in research and supervision

TDHB is Professor of Biological Chemistry, he has been an independent academic since 1991, and has published to date 85 journal publications at the interface of chemistry with biology (of which 42 have named PhD students in TDHB's research group). His area of research is enzymology, in particular enzymes involved in bacterial peptidoglycan biosynthesis, and enzymes involved on bacterial aromatic degradation pathways. He has supervised 22 PhD students to successful completion (+ 1 recently submitted, 1 writing up), with a 96% submission rate within 4 years. All of the PhD students in his group have carried out interdisciplinary PhD projects involving training in chemistry and protein biochemistry. CGD is Professor of Microbiology in the Department of Biological Sciences. His area of research is medical microbiology, in particular the bacterial genetics of high-level penicillin resistance in *Streptococcus pneumoniae* (67 journal publications). At Warwick he has supervised 12 PhD students, 2 MD students and 1 MPhil student to successful completion, with a 94% submission rate within 4 years. DIR is Associate Professor in Structural Biology in the Department of Biological Sciences. His area of research is protein crystallography (46 journal publications). His area of research is enzymology and structural biology, in particular enzymes involved in bacterial peptidoglycan biosynthesis, cell division and bacterial aromatic degradation pathways. He has supervised 3 PhD students to successful completion and has a further 4 students currently registered, with a 100% submission rate within 4 years.

2. Background to research project

TDHB (Chemistry), CGD, and DIR (Biological Sci.) have collaborated since 2002 to study the cytoplasmic and lipid-linked steps of bacterial peptidoglycan biosynthesis, in particular the mechanism of high-level penicillin resistance in *Streptococcus pneumoniae*. This collaboration has led to the reconstitution of the tRNA-dependent MurMN reactions involved in high-level penicillin resistance in *Streptococcus pneumoniae* [1], and the development of the first inhibitor of this class of tRNA-dependent ligase [2]. We have also developed novel fluorescence-based assays for the later steps in peptidoglycan biosynthesis [3,4], which will enable us to screen for novel inhibitors as leads for new antibacterial agents, and have undertaken structural biology on peptidoglycan biosynthetic enzymes [5]. This research has been supported by the Wellcome Trust and MRC, and we have recently obtained a joint grant with Dr. Mike Chappell and Dr. Neil Evans (Engineering) to model quantitatively the peptidoglycan biosynthetic pathway, funded by EPSRC. We have established a Network for Bacterial Cell Wall Assembly (BACWAN), which is funded by MRC, and through which we have established a facility to prepare each of the peptidoglycan biosynthetic intermediates via biocatalysis. TDHB has collaborated since 2001 with Prof. S. Evans (University of Leeds, Physics), to develop physical methods to study the lipid-linked steps of peptidoglycan assembly, using supported lipid bilayers. Lipid intermediates I and II have been deposited into a supported lipid bilayer, and surface plasmon resonance (SPR) has been used to measure the binding of vancomycin and ramoplanin, and the *in vitro* polymerisation to form peptidoglycan [6,7]. TDHB and SDE have a current BBSRC grant to study *in vitro* peptidoglycan biosynthesis using supported lipid bilayers, and a variety of quantitative physical methods.

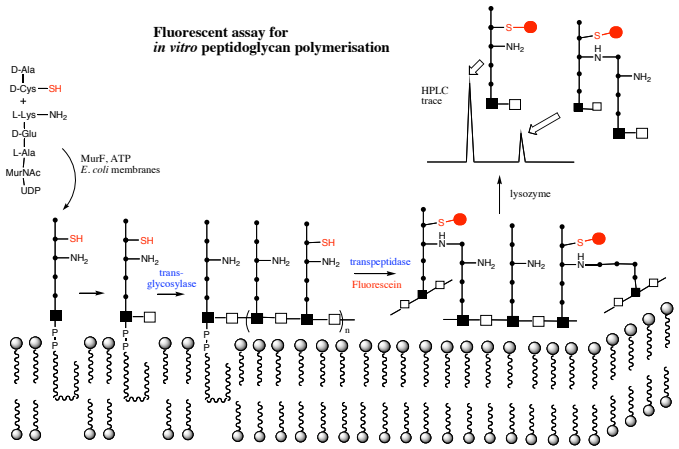
The penicillin binding proteins (PBPs) catalyse the final two steps of bacterial peptidoglycan biosynthesis: namely transglycosylation of lipid intermediate II to form a glycan chain, and transpeptidation of the glycan chains via cross-linking of a lysine residue (or *meso*-diaminopimelic acid in Gram-negative bacteria) on position 3 of a pentapeptide sidechain with the D-alanine residue at position 4 of another pentapeptide [8]. These reactions are very important antibacterial targets, since the transpeptidase activity is the molecular target for the penicillin family of b-lactam antibiotics, and both steps are inhibited by the vancomycin group of glycopeptide antibiotics. Many strains of penicillin-resistant bacteria contain alterations in their PBPs, therefore these proteins are also the site of important antibiotic resistance mechanisms. However, in spite of the large amount of research into b-lactam chemistry, there has been little detailed enzymology carried out on the PBPs, due to a lack of a quantitative assay method to study trannsglycosylation and transpeptidation. In part this is due to the inherent difficulties in preparing the lipid II substrates for PBP-catalysed reactions, but we are now able to prepare lipid intermediate II in mg quantitities [1]. We have developed a new strategy for fluorescent labeling of lipid II [3] by incorporation of an unnatural D-Cys residue in place of D-Ala in the pentapeptide sidechain, which we will use to develop a novel PBP assay. We are therefore in a unique position to develop quantitative methods to study the PBP-catalysed reactions, that could be used to study the enzymology of PBPs from a range of bacterial strains, including PBPs from antibiotic-resistant bacteria.



3. Research project

The aim of the project is to develop new quantitative assays for the transglycosylase and transpeptidase activities of the penicillin-binding proteins, and to use these assays to characterize quantitatively several PBPs from penicillin-sensitive and –resistant strains of *Streptococcus pneumoniae* and *Staphylococcus aureus*. The project will be divided into the following sections:

A) **Development of quantitative assays for transglycosylase (GTase) and transpeptidase (TPase) activities (months 1-12)**. We have recently found that, via incorporation of dipeptide D-Cys-D-Ala into peptidoglycan using *E. coli* membranes, followed by fluorescein labeling, we can generate fluorescent peptidoglycan. Digestion with lysozyme gives two new peaks by HPLC, corresponding to the fluorescent disaccharide-pentapeptide (indicative of transglycosylase activity) and the cross-linked disaccharide-pentapeptide (indicative of transpeptidase activity).



This is a novel, sensitive assay for GTase and TPase activities. This assay will be adapted to use a D-Cys-labelled lipid II substrate, that could be incubated with a single, purified PBP, to measure GTase and TPase activities. The assay will be developed first of all using a bifunctional Class A PBP, *S. pneumoniae* PBP-1B, which has been expressed in DIR's group. Recombinant PBP will be expressed and purified, then the D-Cys-labelled substrate will be prepared, and conditions for reliable assay will be determined. The assay will then be used to kinetically characterize the enzyme (measurement of k_{cat}, K_m parameters).

B) **Reconstitution of *S. pneumoniae* and *S. aureus* PBPs with modified lipid II substrates (months 13-24).** *S. pneumoniae* strains with high level penicllin resistance contain high levels of Ala-Ala and Ser-Ala cross-links, synthesized by the MurMN ligases [1]. We have developed a chemoenzymatic method to prepare modified lipid II-amino acyl conjugates, which we will use to prepare lipid II-Ala-Ala and lipid II-Ser-Ala. We will test these substrates against recombinant PBPs from penicillin-resistant and –sensitive strains of *S. pneumoniae*. Some expression constructs are already available at Warwick, or will be sent to us by Dr. T. Vernet (IBS, Grenoble), other constructs will be cloned as needed during the project. The kinetic parameters will provide important information on the mechanism of high-level penicillin resistance in *S. pneumoniae*. This work will complement an ongoing project in TDHB's research group on polymerization of lipid II-(Gly)5 in *Staphylococcus aureus*.

C) ***In vitro* peptidoglycan synthesis on supported lipid bilayers (months 25-36).** The D-Cys labeled lipid II substrate and recombinant PBPs used above will be used to carry out in vitro polymerization of lipid II in supported lipid bilayers, in collaboration with Prof. S. Evans (Leeds). The architecture of the peptidoglycan layer thuis formed will be studied using confocal microscopy and atomic force microscopy, which will provide exciting new information about spatial and temporal construction of bacterial cell walls, and the roles of each PBP in cell wall synthesis. There are opportunities to reconstitute individual PBP reactions with cell division proteins (FtsW, RodA) that are known to be involved in cell wall synthesis, but their precise biochemical role is uncertain.

- **Originality.** There is currently no use quantitative PBP assay, so the development of a quantitative GTase/TPase assay will be of major interest in this field. Quantitative information regarding the mechanism of high-level penicillin resistance in *S. pneumoniae* would be of considerable interest and impact. Spatial and temporal information about PG synthesis would also be extremely novel in this field.

- **Feasibility**. We have already established the feasibility of the PBP assay using whole E. coli membranes, and the feasibility of supported lipid bilayer work using whole membranes, so it should be quite feasible to extend these methods to individual PBP reactions.

- **Adventure**. This project will develop a completely new PBP assay, and has the potential to apply new surface science methods to PG synthesis.

4. **Scientific strategy**. This project will develop our collaborative approach to studying the lipid-linked steps of bacterial peptidoglycan biosynthesis. The methods developed in this project could be applied to the development of novel high-throughput screens for PBP inhibitors as novel antibacterial agents, and the study of antibiotic resistance mechanisms in a range of important bacterial pathogens.
Relevance to MOAC Scientific Remit. This project fits within the **membrane proteins** applications priority area of MOAC, since PBPs are membrane proteins that catalyse reactions upon lipid-linked substrates. There is also a strong element of **experimental design** within the project, since new assay methods will be developed for the PBPs, and new surface science methods for studying *in vitro* peptidoglycan biosynthesis.

5. **Technical training**. The project will involve training in the following areas: 1) protein expression & purification (done in Chemistry & Biol Sci); 2) preparation of substrates, using solid phase peptide synthesis and biocatalysis (done in Chemistry); 3) supported lipid bilayer work & surface science (done in Chemistry & with collaborator in Leeds); 4) the project may also involve molecular biology cloning work (done in Biol Sci).

6. **Integration of the multi-disciplinary content**. TDHB, CGD & DIR have collaborated since 2002 on peptidoglycan biosynthesis projects, via regular scientific progress meetings. The student will meet on a monthly basis with all 3 supervisors and researchers on related peptidoglycan projects to review progress, and to decide what the next steps in the project will be, and where the next stage of the project will be carried out. Development of the PBP assay (Part A) will be done primarily in Chemistry, with assistance in protein expression from DIR. Cloning & expression of additional PBPs (Part B) will be done in Biol Sci, while preparation of modified lipid II's will be carried out in Chemistry. Supported lipid bilayer work will be carried out in Chemistry, and on visits to our collaborator in Leeds.

7. **Justification of resources required**. Consumables costs of £2K/yr are requested for protein purification cosumables, molecular biology consumables, and peptide synthesis reagents. Additional costs will be met by the groups of TDHB, CGD, and DIR.

8. **Project management**. The principal supervisor will be Prof Tim Bugg (Chemistry), and co-supervisors will be Prof Chris Dowson (Biol Sci) and Dr David Roper (Biol Sci). The student will meet on a monthly basis with all 3 supervisors and researchers on related peptidoglycan projects to review progress. We anticipate that Part A of the project will be done jointly in Chemistry (assay development, kinetics) and Biol Sci (protein expression); Part B of the project will be done jointly in Chemistry (modified lipid II preparation, kinetics) and Biol Sci (cloning & expression); Part C of the project will be carried out mainly in Chemistry, with visits to our collaborator in Leeds.

9. **Plan for what the student will do in the first 3–6 months.** The student will initially become familiar with the existing fluorescent assay, via synthesis of D-Cys-D-Ala, and incubation with *E. coli* membranes & treatment with lysozyme (Chemistry, 1 month). The student will then prepare enzymatically the D-Cys-labelled lipid II analogue, by conversion using *Micrococcus flavus* membranes (Chemistry, 1 month). The student will then learn to purify recombinant *S. pneumoniae* PBP-1B (Biol Sci, 2 months), and will then use the recombinant enzyme to attempt polymerization of the D-Cys-labelled lipid II substrate, followed by fluorescent labeling, and analysis of the peptidoglycan products by HPLC (Chemistry, 2-3 months). During the first 6 months, the student will also receive training in enzyme kinetics (TDHB group), and will attend a training course in radiochemical safety (the project might involve some use of radio-isotopes).

10. References

1. A.J. Lloyd, A.M. Gilbey, A.M. Blewett, G. De Pascale, A. de Zoeiby, R.C. Levesque, A.C. Catherwood, A. Tomasz, T.D.H. Bugg, D.I. Roper, and C.G. Dowson, *J. Biol. Chem*., **283**, 6402-6417 (2008).
2. E. Cressina, A.J. Lloyd, G. De Pascale, D.I. Roper, C.G. Dowson, and T.D.H. Bugg, *Bio-Org. Med. Chem. Lett*., **17**, 4654-4656 (2007).
3. J.A. Schouten, S. Bagga, A.J. Lloyd, G. De Pascale, C.G. Dowson, D.I. Roper, and T.D.H. Bugg, *Molecular Biosystems*, **2**, 484-491 (2006).
4. J.J. Li and T.D.H. Bugg,*J. Chem. Soc. Chem. Commun*., 182-183 (2004).
5. A.M. Blewett, A.J. Lloyd, V. Fülop, C.G. Dowson, T.D.H. Bugg and D.I. Roper , *Acta. Cryst.,* **D60**, 359-361 (2004).
6. M. J. Spencelayh, Y. Cheng, R. J. Bushby, T.D.H. Bugg, J-J. Li, P. J. F. Henderson, J. O'Reilly, S. D. Evans, *Angewandte Chemie Intl. Ed*., **45**, 2111-2116 (2006).
7. C.E. Dodd, B.R.G. Johnson, L.J.C. Jeuken, T.D.H. Bugg, R.J. Bushby, & S.D. Evans, *Bio-Interphases*, **3**, 59-67 (2008).
8. P. Macheboeuf, C. Contreras-Martel, V. Job, O. Dideberg & A. Dessen, *FEMS Microbiol. Rev*., **30**, 673-691 (2006).

Project no. 3

MOAC PhD-project proposal

Project title: **Transcriptional signatures of single-cell lineages in large cell populations**

Supervisor name	Department/email/phone number	Supervisor's advisor
Dr Keith Vance	Biomedical Research Institute, K.W.Vance@warwick.ac.uk, +44 (0) 2476 5 28380	Prof. Georgy Koentges
Dr Till Bretschneider	Warwick Systems Biology Centre, T.Bretschneider@warwick.ac.uk, +44 (0) 2476 1 50252	Prof. Georgy Koentges

Scope of the Project, Multi-Disciplinary Context, Scientific Environment, and MOAC remits

Cis regulatory modules (ReMos) in the non-coding region of the genome regulate the temporal and spatial transcription of the large number of genes that control biological processes. These modules are typically several hundred base pairs in length and contain clusters of transcription factor (TF) binding sites. The aim of this project is to use a tissue culture system to ask fundamental questions about how ReMos function to control cell identity and fate decisions in mesenchymal stem cells.

To this end we have developed a live cell fluorescent reporter based technique in which ReMos are used to drive expression of different fluorescent proteins in order to accurately measure the dynamics of gene expression in real time in individual cells. Measurements are carried out using the Cellomics KineticScan (KSR) machine at its basis and are analysed using custom algorithms. The KSR is an integrated setup containing a confocal microscope with a high resolution CCD camera, a humidified incubator, a liquid handling robot and a computer system allowing imaging of thousands of cells at up to 30 minute intervals over a period of 2-3 days. Importantly, this technique allows the acquisition of high-throughput single cell measurements which eliminate measurement noise inherent in averaging over a heterogeneous population of cells and permits the cell cycle dependent synchronisation of cells *in silico* in order to precisely correlate ReMo activity with cell cycle position. However, a major limitation of current single cell approaches is the identification and tracking of cells both through cell division and in confluent cultures. We will particularly address this problem and develop new image analysis methods to try and improve the percentage of cells that can be automatically followed.

This project will complement a current PhD project in KV's lab in which destabilised reporter proteins are being generated in order to monitor dynamic changes in ReMo activity in mesenchymal stem cells (Danuta Jeziorska) and will form part of a large scale analysis of cis-regulatory function with links to the following SB projects: identification of ReMos using bioinformatics (Sascha Ott and Georgy Koentges) and mathematical modelling of transcription rates (Dan Woodcock, Baerbel Finkenstaedt and David Rand).

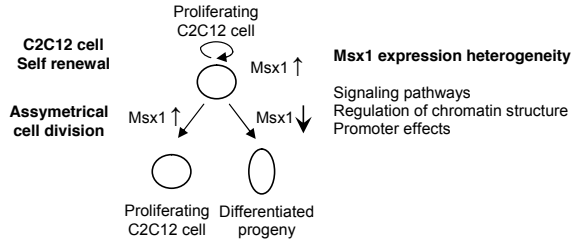
On the image analysis side TB with SO have developed a software interface to import Cellomics data into the image analysis software ImageJ. It has become integral part of the data acquisition pipeline. In preparation of a Systems Biology Miniproject "Classification of Cell Populations in Large Image Data Sets (16/6-5/9/2008, supervisors TB/SO/GK)" we have already developed concrete ideas as to how expand this software to provide thorough statistics of the variability of gene transcription and correlations with cell, and nucleus size. The proposed PhD project will imply the development of novel tracking methods which based on such statistics. and complements TB's project of analysing spatio-temporal patterns of protein distributions in moving cells[*,†].

The candidate is expected to have a strong programming background. The ratio lab/desk will be 30%/70%. The student's budget will be sufficient to cover the consumables required for the experimental part, but still allow for travel expenses to conferences.

The project will be embedded in the following MOAC remits: DNA-protein complexes, Instrument design and development, Experimental design and data Analysis, and Model Building.

Biological Techniques

We will specifically study the mechanisms controlling expression of the homeobox transcription factor Msx1, a master regulator of pluripotency and cell fate in mesenchymal stem cells, using C2C12 cells as a model. In this system Msx1 is expressed in proliferating cells and is presumed to exert its effect through both repressing differentiation and promoting self renewal. Msx1 expression is controlled by four ReMos lying within 5 kb of the transcriptional start site. Stable cell lines containing Msx1 5 kb upstream regulatory region driving expression of a destabilised fluorescent reporter will therefore be generated and used to trace Msx1 ReMo activity through cell division and *in* vitro differentiation on a single cell level. Detailed image analysis generating real time fluorescence response profiles will allow us to measure both the dynamics of Msx1 ReMo activity and to determine the extent of heterogeneity in Msx1 expression levels within individual cells in a genetically identical clonal population. This will enable us to dissect the mechanisms determining noise in Msx1 expression levels and how this affects self renewal and cell fate choices in these cells.



Contributions to heterogeneity in Msx1 expression and possible affects on cell fate choices

Theoretical Methods

The objectives of the theoretical part are a) to develop image analysis routines to reliably track cells and gfp-expression over cell division, b) to extract time courses of nuclear gfp-intensity for cell lineages (at least daughter and grand-daughter cells) and statistically analyse their variation.

-- Evaluation of different segmentation methods

The built-in segmentation procedure of the Cellomics software uses a fixed intensity threshold to identify cells. Because of inevitable photobleaching in long-term experiments this method is prone to fail, because a large number of cells is not recognised toward the end of an experiment. The number of recognised cells is the major limiting factor in the current analysis of the Cellomics experiments and any improvement would be highly appreciated. We have already developed software to access the Cellomics data directly from within ImageJ and successfully experimented with different segmentation methods which recognise shapes as such (Scaling index filters). We will also try filters that are based on structural features (textures).

-- Multi-parametric statistical analysis of cell size, shape and motility, chromatin staining (Hoechst) during cell division and GFP-fluorescence of specific marker proteins

The novel tracking algorithms to be developed will be based on statistical classifiers to discriminate between cells in different developmental stages, i.e. if we know the characteristic temporal dynamics of a nuclear stain in cells undergoing division this information can be built into the tracking algorithm. Other parameters to be included in the analysis are cell size, shape and cell movement, persistence of cell motion, fluorescence intensity, and the temporal dynamics of fluorescence.

-- Development of novel tracking methods

The existing tracking method developed by Sascha Ott is based on a neighbourhood analysis and takes into account distances between cells as well as variations in GFP-fluorescence and nuclear staining. Based on a probability function a cell will be classified as more likely to be identical between one frame and the following if its distance and intensity is closer than that for any other cell in its environment. During cell-division however cells cannot be reliably classified using this method and many cells have to be excluded from the analysis. Together with SO we will develop a new strategy to track cells based on multi-parametric measures (see previous paragraph). A combined forward/backward search to identify cells between successive frames will be employed, which will span multiple time points. One possibility is to expand this into an iterative scheme with several rounds of classification, in which cells that have been tracked with a defined reliability are excluded from the analysis in the next round. Since the neighbourhood information would have to be computed only once, this should increase overall reliability without compromising speed and efficiency too much.

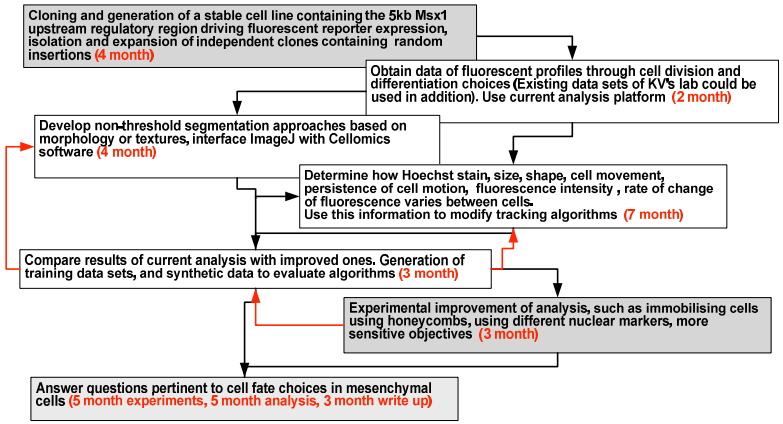
[*] J. Dalous, E. Burghardt, A. Müller-Taubenberger, F. Bruckert, G. Gerisch, T. Bretschneider. Reversal of cell polarity and actin-myosin cytoskeleton reorganization under mechanical and chemical stimulation. Biophys. J., 94(3):1063-1074, 2008.
[†] M. Etzrodt, H. Ishikawa, J. Dalous, A. Müller-Taubenberger, T. Bretschneider, G. Gerisch. Time-resolved Front and Tail Responses to Chemoattractant in Dictyostelium cells. FEBS Letters, 580(28-29):6707-6713, 2006.

-- *Generation of training data sets and synthetic data to evaluate algorithms*
In order to be able to compare different tracking algorithms in an objective manner, training data sets have to be constructed. To avoid tedious full manual tracking of cells we will use automatically tracked sequences and write software which allows inserting tracks manually. Synthetic data of dividing cells based on the statistical distributions determined from experiments will be generated to identify possible tracking problems more easily.

-- *Data analysis*
The data analysis will comprise the alignment of temporal fluorescence profiles (in silico synchronisation), the extraction of lineage trees, and the analysis of variations in expression patterns of daughter and grand-daughter cells. The processed data will be investigated using software by Dan Woodcock to study transcriptional control.

Work Programme

Red arrows denote iteration cycles between experiment and theoretical analysis. Experiments will be performed at BMRI on Gibbet Hill, all dry work in Coventry House.



Keith Vance obtained his PhD at the Cancer Research UK Beatson Laboratories in Glasgow studying cell type specific regulation of papillomavirus transcription. He then obtained a Research Development Fellowship from the Medical Research Council to study the role of T-box factors in the control of transcription and regulation of chromatin structure during melanocyte development and melanoma progression. Following this KV moved to University College London where he initiated his work on making real time gene expression measurements in single stem cells. KV was appointed Senior Research Fellow in Warwick Systems Biology in April 2007 and is currently supervising one PhD student.

Till Bretschneider was appointed Assistant Professor in Warwick Systems Biology in July 2007. He studied Biology at the Ludwig-Maximilians Unervisty in Munich and turned to Computational Biology during his postgraduate studies. During his PhD with Kees Weijer he moved to the University of Dundee and developed computational models for *Dictyostelium* morphogenesis. As a postdoc at the University of Bonn with Wolfgang Alt he continued this work with more abstract mathematical models of mechanical cell-substrate interactions and was awarded a Max-Planck fellowship for his contributions to “Mathematical Models in Biology, Chemistry and Physics”. Joining the group of Günther Gerisch at the Max-Planck-Institut of Biochemistry, Martinsried, in 2001 he expanded his work on quantitative image analysis he had already started in Bonn, primarily working on actin dynamics in *Dictyostelium*. In the past TB has been supervising 2 diploma, 1 bachelor, 1 master and 2 PhD students, which resulted in 5 student co-authorships. In the current academic year TB is supervising one SB and one MOAC miniproject.

Development and Application of Diamond Based Nanopore platforms for the Investigation of Dynamic Events in Ion Channels

Julie Macpherson, Patrick Unwin (Chemistry, Electrochemistry and Interfaces) & Mark Newton (Physics) & Colin Robinson (Biological Sciences)
External Collaborator Prof. Henry White (Chemistry, Utah University, US)

Track Records of Applicants:
Prof. Julie Macpherson is a leading expert in the development of high resolution electrical and electrochemically-based imaging techniques.[11] She has won five awards for her work in this area (including the Marlow medal 2005; McBain medal 2006, Times Higher Education Young Researcher of the Year Award 2005). She has published ~ 100 papers, and during the past 8 years has progressed from University Research Fellow to Professor of Chemistry. **Prof. Patrick Unwin** has held a Chair in Chemistry at the University of Warwick since 1998. He is the author or co-author of almost 200 papers and reviews, has won several international awards for his work and has pioneered new experimental techniques, including the theoretical treatment of coupled interfacial/mass transport problems and applications to a wide range of problems of fundamental and practical importance. Since 1996 the Electrochemistry and Interfaces Group (Macpherson and Unwin) has had 28 PhD students successfully graduate (100% graduation record: see www.warwick.ac.uk/electrochemistry for further details). The group has formed successful collaborations with colleagues in Physics (Newton) and Biology (e.g. Robinson, Pinheiro). **Dr. Mark Newton** is a Reader in Experimental Physics at the University of Warwick who uses a variety of spectroscopic, electrical and microscopic techniques to identify the structure and properties of defects and impurities in diamond. He has a 100 % track record with PhD students successfully graduating ten during his academic career. He is a consultant for the Diamond Trading Company (DTC) and Element Six Ltd and secretary of the UK Diamond Research Committee. Newton, Unwin and Macpherson have several collaborative programmes in place funded by BAE Systems, EPSRC Analytical Studentships and Element Six. **Prof. Colin Robinson** has worked on the transport of proteins into and across chloroplast membranes and his group was first to identify most of the known thylakoid protein transport pathways. He was awarded a BBSRC Professorial Fellowship in 2001 to develop new approaches in this field. He has supervised numerous PhD students with a 100% completion record. **Prof. Henry White** (external collaborator) is head of Chemistry at the University of Utah, USA and assistant editor of *J. Am. Chem. Soc*. He has won 16 prestigious awards during his career history (see http://www.chem.utah.edu/directory/faculty/white.html) with many of his PhD students becoming successful academics. He has had a long standing interest in nanoscale domains and electrochemistry and recently developed the nanopore electrode, based on glass, for the sensing of ion channel events at the single level.

Introduction
Improvements in chemical and biochemical analyses rely on new materials and methods that provide improved sensitivity, reliability and longevity. Increased temporal and spatial resolution enables analyses of biological events at the single molecule level and at very short timescales, providing kinetic, thermodynamic, and chemical dynamic information. To achieve these goals, new methods and materials are continually sought, with a particular emphasis on cost and manufacturability. With this in mind there has been much effort in silicon and silicon nitride microfabrication[1] and glass (to a lesser extent)[2] to make small bioanalytical devices for mass production.

Nanopores and micropore systems are used in many different analytical applications including the detection of single molecules,[3] the detection and sequencing of DNA,[4] particle counting,[5] model drug delivery systems, liquid-liquid analysis, and screening of the interactions of new pharmaceutical drugs with protein ion channels that are found in the human body.[6] The majority of these applications rely on the measurement (often moderation) of an electrical (or electrochemical) signal, be it across an open pore, liquid-liquid interface formed at the mouth of the pore or a bilayer-modified pore. In the majority of cases the size and geometry of the pore and electrical properties of the substrate in which the pore is formed are vitally important factors which control the sensitivity of the device for chemical and biochemical analysis. The chemical stability of the substrate material determines the usefulness of the device in different environments.

One area which is particularly exciting is the measurement of dynamic events associated with protein ion channels. Typically, these are investigated using conductivity measurements in conjunction with model membrane structures suspended over Teflon holes hundred(s) of microns in diameter with inserted ion channels. However this approach suffers from several stability issues, limiting measurements to only hours before the bilayer collapses. To date the fastest, most sensitive substrate for measuring dynamic events associated with protein ion channels is based on the glass nanopore.[7] The bilayers are stable for weeks and it is possible to probe just one ion channel at a time. However, the nanpores are fabricated by hand which is an incredibly time-consuming and laborious process. Furthermore, due to the high capacitance of the bilayer and the support electrical readout of very fast single binding or translocation events is prevented and noise is introduced.

The ideal substrate material in terms of electrical properties is diamond; the resistance and capacitance properties of this material mean that, for the first time, sub-μs measurements should be possible. Diamond is also amenable to machining and etching, so the fabrication of pore shaped structures should be much simpler than using glass based nanopores. Given their robustness, diamond devices are capable of operating under extremes of temperature, voltage and pH conditions. Furthermore, the excellent optical properties means that light activated processes can also be investigated. To this end, diamond is seen as an 'ideal' substrate material to further advances in the field of chemical and biochemical analyses.

Aims
The following developments are proposed which will facilitate major new insights into protein ion channel dynamics and establish a new technology at the physical sciences/life sciences interface:

1) Investigation of chemical functionalisation strategies to produce stable bilayers (e.g. DPPC, DPhPC) on the surface of polished intrinsic diamond (single crystal and polycrystalline). This will be explored using fluorescently labelled lipids in conjunction with confocal microscopy and atomic force microscopy. (Chemistry)
2) Machining and etching strategies to produce micro- to nano-pores in diamond of different geometries (Physics)
3) Finite element modelling of the fabricated pore structures in order to calculate the expected current (conductivity) responses in the presence of open, closed and partially blocked pores (Chemistry and Physics) and to calculate theoretical response times
4) Application of bilayer preparative procedures developed in 1 to the structures made in 2. Initial characterisation and testing and addition of ion channels. Initially commercially available ion channels such a α haemolysin (HL) (well modelled, structure understood, 'open' ion channel) will be incorporated into the membrane and the system optimised such that it is possible to incorporate only one ion channel into the model membrane. In later stages, conductivity measurements will be complemented by combined (simultaneous) conductivity-confocal microscopy measurements to investigate single ion dynamics (Chemistry)
5) Applications: we envisage two main application areas which have huge scope
    a. Building on the α-HL work, we aim to investigate whether it will be possible to identify the base sequence of DNA. It has been suggested that the variation in current, as DNA moves through the pore, might be used to identify the sequence of bases. Due to experimental limitations, high-speed detection of the sequence of bases in single molecules of DNA or RNA to date has not been possible, however using diamond as a substrate enabling the accessing of faster timescales this becomes a reality. This work will be carried out in collaboration

with our colleagues at the University of Utah who have experience with monitoring ssDNA through glass nanopores.

b. We aim to characterise protein conducting channels in the chloroplast thylakoid membrane for the first time. The major protein transporter in this membrane (the Tat system) is driven by the thylakoid proton motive force, and protein transport is known to require large numbers of protons (ca 30k per protein transported). The sensitivity of this technique is such that we will aim to characterise individual translocase complexes by virtue of their proton-pumping activity. By saturating the translocase with high concentrations of substrate, we will be in a position to monitor proton fluxes at very high resolution because protons will be pumped out of the thylakoid at specific locations. By coupling the ion channel device with local microelectrochemical strategies (to generate proton fluxes), we will be able to determine whether the transport of different substrates requires differing quantities of protons. Complementary structural work, notably in-situ AFM coupled to the ion channel devices, will also enable to us to map the positions of the protein translocase complexes on the thylakoid surface. This work will involve thylakoid preparation in Biological sciences, together with the production of recominant Tat substrates (routinely carried out in the Robinson lab).

Obviously there will be flexibility in the workplan, based on the outcomes of 1 – 5 and other application areas which arise during the course of the project.

Training Provided and Facilities

The supervising teams bring multidisciplinary skills which are needed to ensure successful development of new techniques to study dynamic events associated with ion channels, which can be used to study important biological questions. The Electrochemistry and Interfaces Group in Chemistry has state-of-the-art facilities for electrochemical measurements, microfabrication, high resolution AFM imaging, confocal microscopy and electrochemistry, in addition to finite element modelling. The group's facilities are enhanced further through close collaborations with colleagues in Physics for use of electron microscopy facilities. Newton is a world expert in diamond and through existing collaborations with Element Six will be able to supply us (in part) with some of the diamond required for this project. He also has extensive knowledge of machining (in house Physics) and etching technologies and will work with the student in this aspect of the project. In Biology the student will have access to biological preparative procedures for thylakoid and recombinant tat substrate production. The student will receive considerable training across the physical sciences, modelling (using engineering/physics simulation techniques) and the life sciences in tackling several important biological questions (see above).

Justification of Resources Required:

The project is largely practical, although there will be some modelling work. We therefore request £2k pa which will be utilised to purchase diamond substrates (some will be provided free of charge from Element Six but not all); AFM tips; consumables associated with electron microscopy and laser micromachining; lipids and ion channels. It is likely our US collaborator will cover some of the costs associated with visits overseas but we will also seek funds the American Travel Fund.

Project Management: Prof. Julie Macpherson will take responsibility for the Chemistry experimental side. Unwin will supervise the modelling of device responses. Newton will supervise the diamond fabrication work and Robinson the work on translocase activity. Macpherson/White will oversee the DNA studies (at Warwick and at Utah). Macpherson, Newton and Unwin already meet once a week to discuss existing diamond related projects and these meetings will be widened to include the proposed project. For this purpose they will meet with the student together on this frequency. Robinson will participate every month for regular updates, and more frequently when the work involves the thylakoid membrane.

MOAC Project 4

Roles of Different Departments:

The majority of the work is Chemistry led and it is envisaged the student will spend about 60% of his/her time there. 20% will be in Physics developing the fabrication procedures for the diamond ion channels and ~ 20% in Biology.

Initial Work Plan (3 – 6 months):

• Functionalisation of planar diamond surfaces in order to produce both stable bilayer and monolayers (training in confocal and atomic force microscopy)
• Training in laser micromachining – development of a wide array of test structures
• Training in finite element modelling – construction of geometries based on the fabricated test structures
• Initial test experiments for conductivity through machined channels, comparison with modelling

REFERENCES

[1]. Dekker, C. *Solid State Nanpores*, *Nature Nanotechnology*, **2007**, *2*, 209-215

[2]. Fertig, N. *et al*, *Activity of single ion channel proteins detected with a planar microstructure Appl. Phys. Lett.* **2002**, *81*, 4865-4867

[3]. Deamer, D. W., Branton, D. *Characterisation of Nucleic Acids by Nanopore Analysis*, *Acc. Chem. Res.* **2002**, *35*, 817 – 825

[4]. Kasiananowicz, J. J., Branton, D., Deamer, D. W. *Characterisation of Individual Polynucleotide Molecules Using a Membrane Channel*, *Proc. Natl. Acad. Sci.* **1996**, *93*, 13770 – 13773

[5]. Bayley, H., Martin, C. R, *Resistive-Pulse Sensings From Microbes to Molecules Chem. Rev.* **2000**, *100*, 2575-2594

[6]. Shaffer, C., Genetic Engineering and Biotechnology News, **2005**, *25*, 1-3

[7] White R. J. *et al Single ion-channel recordings using glass nanopore membranes*, *J. Am. Chem. Soc.* **2007**, *129*, 11766-11775

Project no.5

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: **"Predicting Membership in Regulatory Protein Complexes"**

Supervisor name	Department/email/phone number	Supervisor's advisor*
Sascha Ott	*Systems Biology Centre, s.ott@warwick.ac.uk, x50258*	*David Wild*
Gyanendra Tripathi	*Clinical Sciences Research Institute, 024 7696 8590, G.Tripathi@warwick.ac.uk*	*Sudhesh Kumar*

Project proposal:

Track record of supervisors in research and supervision

Sascha developed a highly sensitive algorithm for the discovery of evolutionarily conserved regulatory modules from DNA sequence data. Sascha has closely analysed a large number of regulatory modules for phylogenetically conserved binding sites and derived several statistically highly significant hypotheses which are currently being tested experimentally in several labs. His current work involves the integration of various data sources such as microarray data, ChIP-chip data, conserved sequences and putative binding sites into the process of gene network estimation. Sascha is supervising two PhD students who started their projects in October 2007. He has supervised a number of mini-projects in both MOAC and SBDTC in the last two years. As his group currently has three members only (one post-doc + two PhD students), there clearly is capacity for expansion.

Gyan and his group are studying the transcriptional regulation of adipokines such as adiponectin and RBP4. Recently, he found a link between an ATP-dependent RNA helicase and inflammation in obese patients. He is well adept in modern molecular biology techniques and performs ChIP assays routinely, which is an absolute requirement for this study. Gyan is supervising one PhD student who started in October 2006.

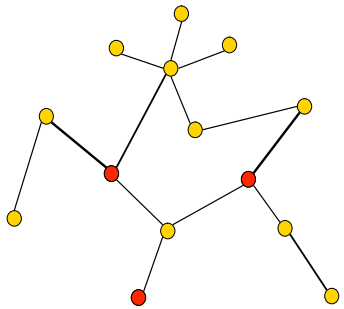


Figure 1. Illustration of the inference task. We are given a network of protein-protein interactions where the thickness of edges represents a confidence score for each protein-protein interaction. We are also given probabilities for each protein to bind to a fixed region of DNA (high likelihoods illustrated as red). The goal is to estimate the likelihood for each protein to be a member of the protein complex assembling on the DNA.

Background to research project

The recruitment of transcription factors (TFs) to regulatory regions is a central event in the transcriptional regulation of genes. A plethora of tools have been developed to predict TF binding based on known sequence motifs. However, instead of directly interacting with the DNA additional proteins can be recruited to regulatory sequences indirectly by forming protein complexes that include DNA-binders (Figure 1). We want to make use of available yeast-2-hybrid databases in order to estimate the likelihood for each protein of a protein-protein network to be a member of complexes that can form on a given region of DNA. You will develop models and methods for this task and apply them to analyse regulatory regions of genes that are differentially regulated between lean and obese patients. You will be able to test selected hypotheses by testing protein presence on the target regions using chromatin immunoprecipitation (ChIP). We will feed positive as well as negative results back into the model to update our estimates in the light of experimental data. This project gives you the opportunity to do research at the important interface of Systems Biology and Medicine. The split between dry and wet work will be roughly **80/20**. Computer programming skills are necessary to pursue this project.

Biological focus

The incidence of obesity and its associated disorders is increasing markedly worldwide. A strong correlation exists between obesity and type 2 diabetes (T2D). Excessive adipose tissue accumulation is a key pathological contributor to the "metabolic syndrome" characterised by insulin resistance and dyslipidemia that leads to T2D and increased risk for cardiovascular diseases. Adipose tissue is no longer considered to be an inert tissue functioning solely as an energy store, but is emerging as an important factor in the regulation of many pathological processes. Several hormones, called adipokines, are synthesised by adipose tissue. These include adiponectin, TNF-α, leptin, resistin etc. Most of these adipokines have been found to be differentially regulated between lean and obese patients. Initially, we will be targeting the regulatory regions of adiponectin gene which we have already mapped using Sascha's software. This study will not only provide a better understanding of the regulation of these genes but also a universal model for studying other systems. This study may also give clues for new therapeutic targets.

Adiponectin (also known as APM1, ACRP30, GBP28, or AdipoQ) regulates insulin sensitivity and energy homeostasis. This gene is located on human chromosome 3q27 and is linked to Type 2 Diabetes (T2D). Several studies have demonstrated that the adiponectin level in plasma is decreased in patients with T2D or obesity in comparison to healthy subjects. Thus, adiponectin may play a role in the pathogenesis of T2D and insulin resistance. Although the physiological effects of adiponectin have been investigated extensively, the molecular mechanisms that regulate the expression of the adiponectin gene are largely unknown. Several factors that regulate adiponectin expression have been reported, such as PPARγ agonist, TNFα, β-adrenergic stimulation and insulin. Our *in silico* analysis of its promoter and upstream region has revealed putative regulatory elements for PPARγ, C/EBP's, NF-Y, Sp1, Sp3, GR, ER, AP2α, YY1, HNF1, NFκB, STAT's and GATA Transcription factors (TFs). The study will also be extended to other genes Resistin, Leptin and TNF-α. They all play important role in regulating insulin sensitivity. Our aims:

(1) *Identification of TF-complexes binding to the regulatory elements:* using ChIP assays we will identify the TF-complexes binding to the regulatory elements

Chromatin Immunoprecipitation (ChIP) assays are used to evaluate the association of proteins with specific DNA regions. The technique involves crosslinking of proteins with DNA, fragmentation and preparation of soluble chromatin followed by immunoprecipitation with an antibody recognizing the protein of interest. The segment of the genome associated with the protein is then identified by PCR amplification of the DNA in the immunoprecipitates. There is a general protocol for ChIP, which is employed by many different labs with minor modifications. We have standardised a protocol for ChubS-7 cells and human adipose tissue in our laboratory.

(2) *Interaction between regulatory proteins:* using GST-fusions and Biacore (This part will only be done if time permits).

For these studies we will be using human preadipocyte cell line ChubS-7 (created by Nestle) which can differentiate into mature adipocytes by changing the media conditions and will be a good representative of normal mature human adipocyte.

The study will also be extended to other genes Resistin, Leptin and TNF-α. They all play important role in regulating insulin sensitivity.

Research project

The strategy on the the biological side of the project is as outlined above. On the theory side we will use existing software (developed by John Reid in collaboration with Sascha) to estimate the probability of DNA binding for each protein that has a known recognition sequence. These will correspond to labels for nodes in Figure 1. We will then employ the approach of (Bader, Chaudhuri et al. 2004) to evaluate the confidence for binary protein-protein interactions based on multiple data sets (thickness of edges). In (Asthana, King et al. 2004) Monte Carlo methods have been employed to sample likely completions of partially known protein complexes from protein-protein networks. We will use an analogous approach by sampling DNA-binders (node probabilities) as well as indirect binders (edge probabilities). This approach

will provide flexibility to estimate probabilities for any feature in the graph and to incorporate additional information. For example, while a single protein may have a large number of potential interactors only a limited subset of these can be bound at any point in time.

Strategic scientific development for Warwick or a research area
This project will facilitate the application of Systems Biology methods to medical research at CSRI. While this project is focused on genes that are differentially regulated in obesity, the methods developed will be generally applicable to any regulatory region of any species for which large-scale yeast-2-hybrid data sets are available.

Technical training
Sascha will provide training in developing new effective methods for the analysis of high-throughput data, Gyan will provide training for lab work.

Integration of the multi-disciplinary content
This project is about analysing public high-throughput data sets in order to shed light on regulatory mechanisms that involve more than a single TF. Experimental results can inform the models such that we can a) assess the reliability of estimations, and b) partially reveal the regulatory complexes forming on a selected set of obesity-associated genes in multiple iterations of the dry-wet cycle.

Justification of resources required
Sudhesh Kumar has committed to contribute £5k per year in addition to the £2k from MOAC. These £7k will provide plenty of consumables for ChIP experiments. All necessary equipment is in place and available for this project.

Project management
We will monitor progress and multi-disciplinary integration in regular joint meetings. Most of the data used in this project is publicly available, we will obtain additional data from ChIP-experiments. By aliquoting sonicated as well as precipitated DNA we will test the binding of multiple TFs to multiple target regions. We expect to test up to 30 TF/target-pairs in one go.

Plan for what the student will do in the first 3–6 months
We will use existing methods for evaluating confidence for binary protein-protein interactions and then start method development by implementing the straightforward sampling approach (see above). We will also do the experimental training so that you can do ChIP independently.

References
Asthana, S., O. D. King, et al. (2004). "Predicting protein complex membership using probabilistic network reliability." Genome Res **14**(6): 1170-5.

Bader, J. S., A. Chaudhuri, et al. (2004). "Gaining confidence in high-throughput protein interaction networks." Nat Biotechnol **22**(1): 78-85.

Project no.6

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: **"Modelling TF Binding Sites and their Combinatorial Effects"**

Supervisor name	Department/email/phone number	Supervisor's advisor
Sascha Ott	*Systems Biology Centre, s.ott@warwick.ac.uk, x50258*	*Jim Beynon*
Jim Beynon	Warwick HRI, jim.beynon@warwick.ac.uk, 024 7657 5141	N/A

Project proposal:

Track record of supervisors in research and supervision

Sascha developed a highly sensitive algorithm for the discovery of evolutionarily conserved regulatory modules from DNA sequence data. Sascha has closely analysed a large number of regulatory modules for phylogenetically conserved binding sites and derived several statistically highly significant hypotheses which are currently being tested experimentally in several labs. His current work involves the integration of various data sources such as microarray data, ChIP-chip data, conserved sequences and putative binding sites into the process of gene network estimation. Sascha is supervising two PhD students who started their projects in October 2007. He has supervised a number of mini-projects in both MOAC and SBDTC in the last two years. As his group currently has three members only (Laura + two PhD students), there clearly is capacity for expansion. Furthermore, Laura will be closely involved in day-to-day supervision.

Jim is PI on the SABR project and has played a major role in the analyses of plant pathogen interactions at the molecular level. His work pioneered the development of the model interaction between the downy mildew parasite *Hyaloperonospora arabidopsis* and Arabidopsis. His lab works on understanding the mechanisms by which pathogens suppress innate immune responses.

Laura Baxter is a senior postdoc on the SABR project working with Sascha on developing the motif search algorithms. She has worked on bacterial genome sequencing, eukaryotic comparative genomics and lead the genome assembly project for *H. arabidopsis*. You will work closely with Laura during your project.

Background to research project

Gene regulatory networks depend on the binding of gene regulatory proteins called transcription factors (TF) to promoters. A new experimental technique called ChIP-seq (Chromatin immune precipitation sequencing) is making it possible to obtain genome-wide data on regions of transcription factor (TF) binding. In this procedure TFs are allowed to bind to their target DNA and are then chemically modified to stabilize the complex. The DNA is sheared and antibodies used to isolate the TFs bound to their target DNA. This DNA is then amplified and sequenced using the next generation sequencing Solexa platform. This generates millions of DNA sequence tags that can be used to define the genomic target regions. The real challenge now is to develop motif finding techniques to identify the exact binding positions and to learn

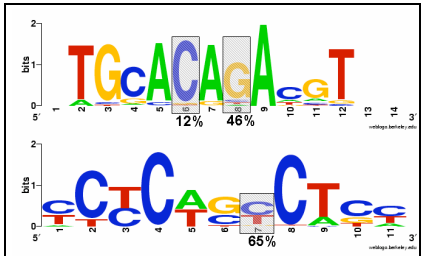


Figure 1. The large number of fragments in ChIP-on-chip data allows fitting of more complex models of TF-binding. Top: an artificial motif containing optional bases was implanted into randomly generated sequences and accurately recovered by our motif-finding software. The height of nucleotide letters represents their conservation among different sites, characters in shaded boxes occur with indicated frequencies. (original frequencies of the two optional bases were 10% and 50%). Bottom: application to real data revealing an optional base within a recognition sequence.

mathematical models of TF binding. These models may reveal combined binding sites of a TF and its co-factor (Loh, Wu et al. 2006) and can be used to identify binding sites in species for which experimental data is not available. Furthermore, such models can explain some of the variation of ChIP fold-changes that is brought about by binding sites of varying affinity (Tanay 2006) and other factors.

The ChIP-seq technology is being established at Warwick and will provide the opportunity to investigate generalised models of TF binding. We will develop methods to find motifs in ChIP-seq data (Figure 1) and mathematical models to explain fold-changes elicited by combinations of TF and co-factor binding sites of varying multiplicity, affinity, and spacing. These models for motifs and motif-combinations will then be used to inform the reconstruction of gene regulatory networks, expand ChIP-seq results to other species, and to propose the design of mutated or artificial promoter sequences for experimental testing. Time permitting we will use combined ChIP-seq data sets from multiple TFs to find statistical links between patterns of co-occurrence of binding sites in promoters and expression clusters observed in microarray data. In the experimental part of the project, we will do ChIP-seq experiments for selected factors. The share between dry and wet will be roughly 80/20. Computer programming skills are necessary to pursue this project.

Biological Focus

The project that you will join has already established key datasets that define normal leaf development and how that is affected under challenge from plant pathogens. We are currently developing new datasets with further pathogens and with the abiotic stresses of drought and high light. Your work will contribute significantly to our understanding of the regulatory networks that underpin plant responses to these environmental stresses. We anticipate that many of the networks are common to all stresses and we wish to identify these core regulatory networks. Your work in this project will help us define the regulatory modules that control these networks.

Research project

You will implement and develop algorithms to analyse Solexa data output to be able to extract the TF binding sites. This will allow you to then map the binding sites to the genome and identify the genes under the TF's regulatory control. These data will be mapped back onto our network models to determine if they confirm or contradict the predictions. These data will also inform significant extensions of the network models by adding genes under the TF's regulatory control that cannot be determined by micro-array analyses. The next level of study will be to define the precise regulatory motif with which the TF interacts. For motif-finding we will use Sascha's implementation of the Equi-Energy-Sampler (Kou 2006) and also try the EM-algorithm for faster convergence. For fitting models of the combinatorial action of binding sites we will use a combination of exhaustive searching and local optimisation similar to the approach chosen in (Tanay 2006). This will allow us to try different types of models without having to make big changes to the optimisation procedures. We will explore fully probabilistic approaches for those types of models that we find to work. For assessing statistical links between combinations of binding factors and microarray data clusters we will employ probabilistic models similar to (Segal, Yelensky et al. 2003; Kundaje, Middendorf et al. 2005) and fit these using EM.

Strategic scientific development for Warwick or a research area

The methods developed in this project will be relevant not just to the SABR projects, but to most of the projects that will test TF binding using ChIP-seq. Therefore, this project will provide useful tools for a larger community in Warwick and internationally.

Technical training

Although this is primarily a dry project using data produced in the SABR project the student will have the opportunity to select a TF and carry out the state of the art ChIP-seq technique. The TF will be cloned under the control of its native promoter, a constitutive high expression promoter and an inducible promoter. The reason for this is that expression levels of the TF can alter binding of the TF to target DNA and it is important to be able to eliminate artifacts in the analyses. You will transform plants and analyse that they are producing the modified TF. We will then either use antibodies to the native protein or to protein tags fused to the native TF to carry out a ChIP experiment. Amplified DNA will be adapted for Solexa analysis and run on the new Genome Analyser G2. This will generate some 10Gb of sequence data of 36 base pair

tags as paired end reads (this means you will be able to define the exact region to which the TF is bound). You will be supported by other members of the SABR team (postdocs and technicians) during this project.

Integration of the multi-disciplinary content

A major new systems biology research project called "Elucidating Signalling Networks in Plant Stress Responses", funded under the BBSRC "Systems Approaches to Biological Research" (SABR) initiative, has just started at Warwick. In this project we are creating network models that describe the transcriptional responses of the plant to environmental stress induced by biotic and abiotic stresses. Using clustering algorithms and state space modelling techniques on high resolution microarray datasets we have already built preliminary coarse-grain gene regulatory network models. Several of the nodal genes in these models encode TFs. We would predict that the TF will either interact directly with the promoters of downstream genes or via intermediate steps not visible in the data. We will test the predictions of the models in this work by carrying out ChIP-Seq experiments with one or two key TFs and identify and model their DNA binding sites. The models you will develop in this project will shed light on how binding affinity, binding site multiplicity, and spacing rules affect the ability of regulatory sequences to recruit TFs. These will inform the coarse-grain network reconstruction by identifying evolutionarily conserved binding sites, and will also provide clues for the fine-tuning of existing network models by providing explanations for early vs late responses etc. Furthermore, we can use these models to predict the effect of mutations in promoters.

Project management

We will monitor progress and multi-disciplinary integration in regular joint meetings. The ChIP-Seq technology has already been established in a number of labs and HRI are working towards running ChIP-Seq in Warwick by autumn. In case of delays to this schedule we will be able to advance the theory side of this project using both publicly available data sets and synthetic data in the meantime. While this project will have a broad interface to above-mentioned SABR-activities, the theory part will be exclusive to this PhD project.

Plan for what the student will do in the first 3–6 months

The focus in the first half year will be to a) become able to analyse single ChIP-Seq data sets and b) start learning ChIP. This will include further development of an existing motif-finder and a first approach to model the effect of multiple binding sites as described above.

References

Kou, S. C. a. Z., Q. and Wong, W. H. (2006). "Equi-energy sampler with applications in statistical inference and statistical mechanics." Ann Statist **34**(4): 1581-619.

Kundaje, A., M. Middendorf, et al. (2005). "Combining sequence and time series expression data to learn transcriptional modules." IEEE/ACM Trans Comput Biol Bioinform **2**(3): 194-202.

Loh, Y. H., Q. Wu, et al. (2006). "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells." Nat Genet **38**(4): 431-40.

Segal, E., R. Yelensky, et al. (2003). "Genome-wide discovery of transcriptional modules from DNA sequence and gene expression." Bioinformatics **19 Suppl 1**: i273-82.

Tanay, A. (2006). "Extensive low-affinity transcriptional interactions in the yeast genome." Genome Res **16**(8): 962-72.

Project no. 7

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: Rigidity dilution analysis as a method of studying prion protein folding

Supervisor name	Department/email/phone number	Supervisor's advisor*
Rudolf A Roemer	Physics/r.roemer@warwick.ac.uk/74328	
Teresa Pinheiro	Bioligy/T.Pinheiro@warwick.ac.uk/28364	

1. Track record of supervisors in research and supervision

Dr. Teresa Pinheiro is an Associate Professor in Biological Science where she leads the group on Protein Folding and Misfolding in Human Diseases such as Alzheimer's, Parkinson and prion diseases. She has crucial expertise on protein-lipid interactions and protein folding using a complimentary array of interdisciplinary biophysical methods. In the last 10 years she has focused on the studies of misfoldings in connection with neuro-degenerative disease and published 15 papers on that particular topic. She has supervised 6 PhD students and several PDRA and research technicians. Her research has been funded by the BBSRC, EPSRC, Welcome Trust and the MRC.

PhD/MSc supervision at Warwick: 4 successful PhD theses, 3 MOAC Msc theses

Dr. Rudolf A Roemer is Associate Professor in Physics and holds a dual appointment with the Department of Physics and the Centre for Scientific Computing at the University of Warwick since 2002. His research has focused on themes of condensed matter theory with a broad range of topics and methods spanning from the mathematical physics of exactly solvable interacting quantum many-body systems to the applications of computational localization theory for electronic transport in DNA. He has supervised 6 doctoral and 3 diploma/MSc theses and his group currently consists of 4 Ph.D. students and 1 PDRA. He is author of more than 100 scientific publications since 1993 and his research is currently funded by EPSRC, the Leverhulme Trust and BBSRC. Rudolf currently serves as Director of Warwick's Centre for Scientific Computing.

PhD/MSc supervision at Warwick: 2 successful PhD theses, 2 successful MSc theses (since 2002)

Dr. Stephen Wells is currently PDRA in Dr. Roemer's group, working on theoretical studies of electronic transport in DNA. Before coming to Warwick in late 2007, he spent three years (2004-2007) at Arizona State University in the biophysics group of Prof. Mike Thorpe, where the FIRST3.0 rigidity analysis code was developed. During this time he implemented geometric simulation for protein structures as the FRODA algorithm, a module of the FIRST code. He has several publications on the algorithm and its application to biophysical problems, including protein complex formation and structural fitting to low-resolution data. We envisage that he will be a key asset to the project. Stephen obtained his PhD in the Earth Sciences department at Cambridge University, where he developed the geometric simulation approach as an extension of the rigid-unit model for framework aluminosilicate minerals, and he has worked in mineral physics at the Royal Institution in London.

2. Research project

a. Originality

The overall objective of the PhD is to use rigidity analysis and geometric simulation to examine the structure-function relationship in *prion* proteins. Prion proteins (PrP) have been shown to undergo upon folding a transition to the pathogenic scrapie isoform. A proper understanding of this transition will be of great medical importance not only with regard to BSE/scrapie/CJD, the diseases associated with prion protein misfolding, but also with regard to general mechanisms of protein misfolding and fibril formation. The computational method we propose to use in this research is very rapid and computationally inexpensive and is therefore more easily applied than molecular dynamics (MD) in two important areas- interpretation of experimental results, and hypothesis testing.

b. Feasibility

We can commonly obtain data describing the behaviour of a protein on a relatively coarse scale. For example a structural or folding transition can be observed using circular dichroism to obtain the amount of alpha and beta structure, or FRET can be used to determine that two domains of a protein have moved together or apart. Simulations are then required to interpret these data in terms of changes to the protein structure on the atomic scale. However, significant transitions such as those during folding of functional motion frequently occur on timescales larger than nanoseconds and hence currently inaccessible to MD. Using geometric simulation we can explore the full range of flexible motion accessible to the protein, and simulate pathways for expansion, contraction and domain motion.

Hence the present proposals aims to understand the behaviour of a protein using atomic-detail models and rigidity-based coarse-graining, giving rise to qualitative (does there exist a viable/permissible (mis-)folding pathway?) and quantitative (at what level of energy does which part of the prion protein become flexible and misfold?) understanding of PrP rigidity and folding.

c. Adventure

These rigidity-based techniques, while showing great promise, are still fairly new and have as yet been applied to only a small number of proteins. Several important systematic questions remain to be answered. One of these concerns the nature of the rigid-to-flexible transition in protein structures. Glassy networks such as chalcogenide glasses can display either first-order or second-order rigidity transition behaviour depending on the amount of variability in local structure (Sartbaeva et al. 2007). It is currently not known whether protein rigidity transitions display the same variation. A second major question is the robustness of the patterns of rigidity and flexibility in a protein as the structure varies, for example when a protein is crystallised under different conditions or as a function of evolutionary changes in the protein's sequence and structure. In a current systems biology project we have begun to examine these questions by examining, for selected proteins, a range of structures drawn from multiple species and conditions of crystallisation. For this particular PhD project the value of the software will be in its capacity for rapid hypothesis testing.

By hypothesis testing, we mean the answering of if-then questions about the influence of the protein's structure on its flexible motion. By performing rigidity analysis and geometric simulation of a structure with and without the presence of a particular bond or interaction, we will determine whether that interaction is significant to the protein's stability- that is, whether it affects the structure's rigidity- and whether it affects the protein's flexible motion. This is directly relevant to mutagenesis experiments as it will suggest residues that are the most suitable targets and predict the effect of, for example, the elimination of particular disulphide bonds and hydrophobic interactions. Again, the low computational cost and rapidity of the simulations will allow us to explore many possibilities in parallel and select the best candidates for experimental investigation.

3. Scientific strategy

The “pebble game”, the idea at the heart of the present approach, is an algorithm based on matching degrees of freedom against constraints, which can be used to determine what part of a structure is rigid and which part is floppy. It has been used to study molecular frameworks including the crystal structure of proteins. It is, however, a great simplification when compared to the standard molecular-mechanics/dynamics approaches. Rigidity analysis provides a natural multi-scale coarse-graining for simulations, particularly using the technique of geometric simulation (Wells et al. 2005). This has the potential to explore the flexible motions of a protein structure far more rapidly than is possible using computationally expensive techniques such as molecular dynamics. The pattern of rigidity and flexibility observed when analysing a protein structure depends on the set of hydrogen bonds and hydrophobic interactions that are included in the analysis.

Previous research suggests that prion protein misfolding may be connected to association of the protein with lipid membranes (Sanghera and Pinheiro, 2002; Kazlauskaite et al. 2003; Pinheiro, 2008). Hypotheses regarding the effect of this association can be tested using rigidity analysis, as the hydrophobic interactions that are important to protein stability in aqueous solution will not be present in a lipid environment. Several NMR structures are available as input for this analysis already from previous work and easily accessible in the Protein Data Bank. Rigidity analyses and flexible motion simulations with the elimination of hydrophobic interactions from portions or facets of the protein will allow us to explore the effects of lipid association on the protein's stability and conformational transitions. The simulations will also allow us to explore the effects of known pathogenic point mutations, by examining the effect on rigidity and flexibility of the interactions of the mutant residue. Ongoing

work in collaboration with the NMR group in Birmingham will also input in this project by providing further structural information on PrP in association with lipid membrane.

4. Technical training

Rigidity analysis will be performed using the pebble-game algorithm as implemented in the protein analysis software FIRST (available from flexweb.asu.edu). Geometric simulation will be performed using the FRODA module, which is part of the FIRST code. During the project we will develop a truly systematic method for applying these techniques to interpret and guide experimental studies on protein structure and function. Initial studies during a Systems Biology miniproject have shown that this approach is feasible and can be performed on existing computational hardware at the Centre for Scientific Computing of which the theoretical supervisor, RA Roemer, is a member. Dr. Wells is an experienced user and developer of these rigidity-based approaches and will provide direct training to the student.

In order to pursue these theoretical techniques with confidence, it is expected that the student attends at least the lecture models CY901 and CY902 of the CSC MSc in Scientific Computing. These are run in week-long intensive courses as well as in term 1 (CY901), term 2 (CY902) and term 3 (CY901), giving much flexibility to the student.

5. Integration of the multi-disciplinary content (this may include a track record of collaboration between supervisors and an indication of how the student and supervisors will interact)

The greatest value will be obtained from these techniques if used in close collaboration with experimental investigations of particular proteins of biological and medical significance. The aim will be to achieve a synergy between the experimental and theoretical work, in which the simulations will aid in the interpretation of experimental data and also suggest the most fruitful avenues to explore in the next round of experimental work. We are therefore forming a collaboration between the theoretical, modeling group of Dr Roemer and the experimental laboratory group of Dr Pinheiro in Biological Sciences, whose group is currently pursuing experimental research on the prion protein (PrP) and its transition to the pathogenic scrapie isoform. It is envisaged that the student spends at least 2 3-months periods in Dr Pinheiro's group on the preparation of the PrP's and their structural characterization by fluorescence, circular dichroism and electron microscopy. These PrP's they will then also model computationally and compare to the experimental findings. Generally, we expect that 25-30% is the project involves work in the Pinheiro group.

6. Justification of resources required

The student will be exposed to the production, purification and characterization of the folded states of prion intermediates. This will include spectroscopic techniques such as circular dichroism, infrared spectroscopy and fluorescence. The student will work in tandem with other PhD students and PDRA working on ongoing projects in the Pinheiro lab on folding of prion protein and conformational changes of PrP in lipid membranes. This type of work is normally costed at 10 to 15k GBP per year. As the student is expected to work up to 6 months in the laboratory, we are budgeting the proposal at 5000GBP for the wet lab component.

Initial studies during a Systems Biology miniproject have shown that this approach is feasible and can be performed on existing computational hardware at the Centre for Scientific Computing of which the theoretical supervisor, RA Roemer, is a member. Hence, we would only need a Linux desktop (1500GBP) as input in year 1 of the project.

7. Project management

The overall project management responsibility will lie with the main supervisor, Dr. Roemer. He will oversee the direction of research and the training requirements for the student. However, the day-to-day research activities will be directed by Dr. Wells, who is a well-known expert in the field. Regular monthly progress meetings with Dr. Pinheiro shall be the main avenue of interaction with the biology team in the initial 6 months phase of the project. Afterwards, the student is expected to spend at least 2 three-monthly periods of wet-lab work in Dr. Pinheiro's laboratory and during that period Dr. Pinheiro will be responsible for the student's progress and welfare. We also envisage that in the later stages of the project, the interaction with the biology team will be enhanced to regular weekly or biweekly meetings. During the PhD period, the student will be expected to participate in the activities of the Centre for Scientific Computing such as, e.g. the weekly seminar series CSC@Lunch and the postgraduate

seminars. Additionally, the student is expected to maintain a visibility in both Physics and Biology activities as well as, most importantly, MOAC.

8. Plan for what the student will do in the first 3–6 months

The student will become familiar with the use of the rigidity analysis and geometric simulation software, FIRST/FRODA, under the direct guidance of Dr. Wells. This will enable the student to perform rigidity analyses on the NMR-derived models of prion protein structures. The student will learn how to edit the set of hydrogen bonds and hydrophobic interactions included in the rigidity analysis, so as to simulate the influence of non-polar (lipid) environments on the protein's structural stability, and to perform geometric simulations of the large-scale flexible motions accessible by the protein. These are the key theoretical/computational skills that will enable the student to become a self-motivated and productive researcher through the remainder of the project. Through literature review the student will also become familiar with Prof. Pinheiro's research and the current state of knowledge regarding prion protein structure and the role of lipid environments in the pathological misfolding transition. By the end of this initial period, the student will be capable of using the software to model and interpret experimental data on the prion protein structure and misfolding transition, and to generate hypotheses susceptible to experimental test by Prof. Pinheiro's group, especially during the student's wet-lab component.

9. References

DC Jenkins, ID Sylvester, TJT Pinheiro (2008). The elusive intermediate on the folding pathway of the prion protein. *FEBS Journal* **275, 1323-1335.**

MR Hicks, AK Rullay, AC Gill, ID Sylvester, IK Bath, DH Crout, TJT Pinheiro (2006). Synthesis and structural characterisation of a mimetic membrane-anchored prion protein. *FEBS J.* **273**, 1285-1299.

J Kazlauskaite, N Sanghera, I Sylvester, C Vénien-Bryan, TJT Pinheiro (2003). Structural changes of the prion protein in lipid membranes leading to aggregation and fibrillization. *Biochemistry* **42**, 10010-10023.

D.J. Jacobs and M.F. Thorpe (1995) Generic Rigidity Percolation: The Pebble Game. *Phys. Rev. Letts.*, **75**, 4051-4054.

D.J. Jacobs, A.J. Rader, M.F. Thorpe, and L.A. Kuhn (2001) Protein Flexibility Predictions using Graph Theory. *Proteins*, **44**, 150-165.

A.J. Rader, B.M. Hespenheide, L.A. Kuhn and M.F. Thorpe (2002) Protein Unfolding: Rigidity Lost. *Proc. Natl. Acad Sci*., **99**, 3540-3545.

B.M. Hespenheide, A.J. Rader, M.F. Thorpe and L.A. Kuhn (2002) Identifying Protein Folding Cores: Observing the Evolution of Rigid and Flexible Regions During Unfolding. *J. Mol. Graph. & Model.*, **21**, 195-207

S. Wells, S. Menor, B.M. Hespenheide and M.F. Thorpe (2005) Constrained geometric simulation of the diffusive motions in proteins. *Phys. Bio.*, **2**, S127-S136

MOAC project 10

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: Statistical mechanics modelling of *MyoD* gene transcription & regulatory module interactions

Supervisor name	Department/email/phone number	Supervisor's advisor*
Dr Mario Nicodemi	Physics/Complexity, m.nicodemi@warwick.ac.uk	
Prof Georgy Koentges	WSB, g.koentges@warwick.ac.uk	

Project proposal:

1.Track record of supervisors in research and supervision

Dr Mario Nicodemi has a joint appointment, as Associate Professor, between the Physics Dept. and Complexity Science Centre at the University of Warwick. A PhD in Theoretical Physics 1997, postdoc in ESPCI (Paris) and Imperial College (London), he has been working on the Statistical Mechanics of Complex Systems in Physics and, more recently, on its applications to Biology. He published about 90 papers in international scientific journals (including Nature, PRL, Genetics, PLoS Comp.Bio.) with about 1200 citations in the ISI Web of Science. He supervised 5 PhDs and several postdocs who published in top Physics Journals, e.g., Nature Materials or Phys. Rev. Lett. Please refer to his web page for some titles and details http://www2.warwick.ac.uk/fac/sci/physics/staff/academic/nicodemi .

Prof Georgy Koentges, appointed in 2007 as experimental co-director of WSB is an international expert in single cell transcriptome profiling and dynamic measurements of gene activity in the area of vertebrate developmental biology and evolution. After studies in comparative anatomy, evolution and philosophy in Tuebingen and Freiburg, he was trained as a molecular embryologist by Prof A Lumsden, FRS, was awarded the T.H.Huxley medal for the best UK PhD thesis in the life sciences, then worked with Prof Dulac and Nobel-prize winner Richard Axel on deciphering the molecular anatomy of the olfactory system through gene-targeting approaches, published in *Cell*, *Science* and *Neuron*. Since starting his own lab in 2001 (appointed as Senior Lecturer at UCL) he published in *Science*, *Nature* etc, developed novel strategies to expression profile single cells from tissues and has been addressing some fundamental questions pertaining to the molecular and cellular evolution of shapes (recently outlined in Koentges, *Nature*, Feb 8th, 2008). As part of this goal he has more recently turned his attention to understanding the logic of cis-regulatory coding in the genome as measured in single cells and has developed with Keith Vance, Sascha Ott and others at WSB an imaging pipeline for single cell tracking of gene activity patterns. His work is conducted with a number of theory collaborators on campus and in the US and is funded over the next four years by two Programme grants of the Wellcome Trust as well as a Human Frontiers science programme grant with collaborators at Stanford and Israel (ranked 6/780 international grant applications). He has supervised a number of PhD students (full-time and part-time) resulting in publications in *Nature* and other journals. All PhDs were completed in the time allocated/available through studentship/grant funding (for part-time students). Current work of a (part-time) PhD student is being written up for *Nature* and is the biological basis of the current proposal that is dominant on the theoretical side (about 70% theory/Physics and 30% molecular experimentation at WSB).

2.Background to research project

The spatiotemporal dynamics of gene action is controlled by regulatory modules (ReMos or CRMs in this text) that interaction with each other and/or the promoter to change the probability of PolymeraseII-HEC promoter docking and escape. Discovery of Remos and their measurement has been a significant challenge in the field that the Koentges lab has recently overcome and that has become a core activity of the Warwick Systems Biology Center. In the last few years we have been measuring combinatorial actions of regulatory modules in massively parallel single cell assays both in time, using the Cellomics KSR as well as in snapshot measurements of the steady-state. The present project will investigate which regulatory module interactions are taking place during this steady-state, it will take advantage of experimental and predicted data, generate a physical model of the interaction logic implemented by these ReMos and make prediction on which proteins involved in interactions shall be manipulated to what effect.

3.Research project – unique strengths

The present proposal aims at bringing together expertise from statistical mechanics/physics (Nicodemi) with genomic systems biology (Koentges), interested in understanding the underlying mechanisms of transcriptional dynamics that the Koentges team is measuring in single cells. The focus of the model will be to understand the logic of combinatorial action, the question whether one can attribute logic functions not just to single factors binding on regulatory regions but on entire regulatory modules. Preliminary analyses the PIs conducted have shown that this is indeed the case. Available measurement platforms for gene-regulatory processes are unique in the academic UK landscape. As far as we are aware, explicit statistical mechanics modeling of these processes in combination is nationally unique too. The project is predominantly theoretical (70%) as it aims at generating **a formal statistical mechanics model of the transcriptional process in which DNA looping, the formation of protein-bridges is explicitly contained**, thus experimental work will be combined with that of Mario Nicodemi on the statistical mechanics of regulatory choices.

4.**Scientific strategy**

Background corrected single cell transcriptional data provide a rich source of information regarding how the interplay between regulatory modules drives MyoD expression. The Koentges lab has already generated several dozens of FACS single cell data sets over the past 1.5 yrs which have measured in every single cell (of 50,000 per experiment) 1. GFP under the control of particular combinations of regulatory modules, 2. Forward Scatter, 3. Side scatter and 4. MyoD protein. This is a rich and internationally unique resource for modeling, background estimation, estimations of noise and transcriptional signal of MyoD transcriptional maintenance. The level of GFP expression recorded in presence of a given ReMo construct is depending on the promoting efficacy of the construct and its related physical binding properties to promoter regions. Thus, we can model the expression level associated with a construct by a thermodynamic model describing ReMos' effective physical interactions. That can be cross validated by comparison with 3C data produced in the lab. We will use methods from computational statistics to make inferences regarding the thermodynamic parameters of ReMos' interactions by using single cell data. Typical inference approaches will be used, including non linear regression and Monte Carlo estimates of variance. (iii) By the above technique a scenario of the physical interrelation of ReMos can be derived. At a later stage we plan to investigate coarse grained, but well established, polymer physics models (Self-Avoiding-Walk polymers with binding sites) to describe the system. , By advanced Monte Carlo simulations the dynamics and thermodynamics of the ReMos can be inferred and the actual network of interactions derived. Thus, new predictions can be obtained to be tested against 3C DNA experiments. (iv) Finally, a further development would be to probe, at a more 'microscopic' level, the interplay between the specific transcription factor binding sites found within ReMos. Once we have identified a small number of the most important interactions between individual regulatory modules, we will explore the details of the system behaviour by characterizing the interplay between transcription factor binding sies (TFBSs) on those modules. This should provide a scenario explaining at a finer scale the ReMos' interaction phenomena described before.

5. Feasibility, theory and technical experimental training

Year1. The student should produce a scenario of ReMos' regulatory interactions by computational statistics inference methods applied to the single cell expression dataset, which were previously obtained in the lab. Statistical inference is based on a model already developed by the PI's (see project description). b) After an early scenario of the physical interrelation of ReMos is derived, as described in (sec.a), that information will be included in a coarse grained polymer physics model of the system (Self-Avoiding-Walk polymers with binding sites). This step will be based on the previous experience of the PI's (see Nicodemi et al. Genetics 179, 717 (2008)). By Monte Carlo simulations the dynamics and thermodynamics of the ReMos can be inferred and the actual network of interactions derived. Interestingly, such a network can be cross-checked against current 3C DNA experiments available from the Koentges lab.
c) Finally, an interesting development would be to apply the experience gained at point b) to describe the system at a finer, 'microscopic', level by a similar model including the details of the specific transcription factor binding sites found within ReMos. This should provide a scenario explaining at a high resolution scale the ReMos' interaction phenomena with currently known transcription factors. To this end the Koentges lab has already generated very significant bioinformatics data sets of regulatory modules and the factors that are a. likely bound, b. berified by ChIP and on single cell microarray measurements and c. are part of known protein-protein interaction networks. This gives a detailed preparation for a modeling project of relevant parameters. d. The binding energies of the majority of DNA-binding domains pertinent to this study are known, the physics challenge will be to simulate the different energy states of the entire regulatory module system that could account for the looping of DNA and the synergistic interactions already found on a more coarse-grained level. **a),b) and c)** were the datasets already given to the students in **MOAC 927 module** for generating a coarse model of the network of the inter regulatory module topologies to which Matt made seminal contributions. The present project continues this line of scientific enquiry towards a realistic model (or sets of equiprobable models) of the underlying process.
Year2: On the experimental side statistical mechanics sensitivity and interaction parameters in the system inform which experimental manipulations will be undertaken. Matt will be determining the choice of which individual (and combinations of) factor(s) to manipulate as a result of the model. To this end he will be ordering single siRNAs for direct experiments and later on constructing libraries of siRNA hairpins under the U6 promoter in the Koentges lab. She has already generated a few of these for us that are immediately available for study. Matt will learn design principles and the details of this process so as to become independent in DNA design. He will learn tissue culture and the use of the FACS machine for the necessary single cell assays. Calibration experiments are well established to introduce him into the successful running of the machine.
Year 3: Development of more fine-grained models, testing of models with siRNA libraries constructed by Matt (or commercially obtained), write-up.

6.Integration of the multi-disciplinary content (this may include a track record of collaboration between supervisors and an indication of how the student and supervisors will interact) **see above**.
7.Justification of resources required. We request the full amount allocated to the student. £500 for the theory/computing, the remainder for the experiments, as siRNAs and molecular reagents are expensive. Prof Koentges commits 10K/year towards the implied experiments from his own research funds.

8.Project management
This project was written jointly between **Matt Bano**, Mario Nicodemi and Georgy Koentges, after Matt contacted the supervisors to ask for interdisciplinary projects that could emerge from the topics of the CH 927 MOAC course module. Expected time expenditure is about 70% theory/Physics and 30% molecular experimentation at WSB. Student will meet with both supervisors weekly and will be part of all group meetings and activities. It is planned that Matt will spend most of his time in **the first year** on the coarse-grain and fine-grain model and will learn how to do tissue culture at the end of the first year. In the **second year**, he will develop the models further, depending on FACS experiments in which individual TFs are modified. In the **third year**, the more fine-grained model and larger simulations will be undertaken and the data will be written up for publication and PhD thesis.

9.Plan for what the student will do in the first 3–6 months. See above. Matt will start with refining the existing background correction for the single cell data sets, will then explicitly model the combinatorics of regulatory modules, based on existing FACS data, leading to a first coarse-grained model in the first 6 months. Further strategy is mentioned above.

We hope that this interdisciplinary project in the area bridging gene-regulatory biophysics, systems Biology and molecular biology will find the MOAC panel's approval.

10.References
M. Nicodemi, B. Panning, A. Prisco, "A thermodynamic switch for chromosome colocalization", ***Genetics*** *179* in press (2008).
M. Nicodemi and A. Prisco, "Self-assembly and DNA binding of the blocking factor in X Chromosome Inactivation", ***PLoS Comp. Bio.*** *3*, e210 (2007).
M. **Nicodemi** and A. Prisco, "A Symmetry Breaking Model for X Chromosome Inactivation", ***Phys. Rev. Lett.*** *98*, 108104 (2007).

Matt Bano
Mario Nicodemi
Georgy Koentges

Project no. 11

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: **Synthesis and biological study of E-NTPDase inhibitors**

Supervisor name	Department/email/phone number	Supervisor's advisor*
Dr Martin Lochner	Chemistry, m.lochner@warwick.ac.uk, x50170	Prof Alison Rodger
Prof Nicholas Dale	Biological Sciences, n.e.dale@warwick.ac.uk, x23729	

Project proposal:

1. Track record of supervisors in research and supervision

Dr Martin Lochner (ML): ML was recently appointed (January 2006) as Assistant Professor of Chemical Biology in the Department of Chemistry at the University of Warwick. He obtained a Dipl. Chem. from the University of Zürich (Switzerland) in 1998 and his PhD studies were completed under the guidance of Prof. Wolf-D. Woggon at the University of Basel (Switzerland) in 2003. ML then won a postdoctoral research fellowship from the Swiss National Science Foundation (SNSF) which he took up with Prof. Ian Paterson FRS in the Department of Chemistry at the University of Cambridge. This was followed by an advanced research fellowship from the SNSF which was taken up in the group of Dr. Sarah C. R. Lummis in the Department of Biochemistry at the University of Cambridge. He has thus established a range of biological research expertise as well as his core synthetic chemistry expertise. ML has published 7 papers[ML1-7] in peer reviewed journals and has presented his work at various international conferences. ML has recently won a first grant from EPSRC (EP/E042139/1) which will investigate the site-specific chemical modification of ligand-gated ion channels.

ML is currently supervising a PhD student (start Oct 2006) which is funded by EPSRC (DTA) and has supervised two MChem students in the past. He is also supervising a postdoc which started in February 2008. ML has successfully guided a MOAC student (James Stephenson) through his mini-project. ML has submitted MOAC mini-projects in every round and a PhD project in 2006.

Prof Nicholas Dale (ND): has examined the role of purinergic signalling in the nervous system for more than 10 years. ND's initial studies examined how ATP and adenosine controlled motor pattern generation in the spinal cord of the *Xenopus* tadpole.[ND1-3] These studies collectively comprise the most complete examination of the role of ATP and adenosine in motor systems and motor pattern generation and shows that ATP and adenosine comprise a complex feedback system that controls centrally programmed fatigue of rhythmic motor circuits such that in the absence of sensory feedback the output slows and weakens over time, before spontaneously stopping. Interestingly, the E-NTPDases, that convert ATP to adenosine, constitute a key locus of control for the dynamics of the centrally programmed fatigue.[ND1,2,4]

During this work ND invented the first biosensor for adenosine[ND2] and this opened the way for collaborations with a large number of groups on other aspects of purinergic signalling. Aided by subsequent invention of a biosensor for ATP,[ND5] ND has studied the roles of ATP and adenosine in brainstem cardiorespiratory reflexes,[ND6] brainstem chemoreception[ND7,8] for O_2 and CO_2, the neuroprotective function of adenosine in the hippocampus,[ND9-12] the release and role of ATP during development of the retina,[ND13] and activity-dependent release of adenosine in the cerebellum.[ND14,15] In all of these studies the use of the biosensors to examine ATP or adenosine signalling in real-time in conjunction with conventional electrophysiological measures of neural activity has been essential in giving new insight into underlying mechanisms.

In parallel with this work, ND continues to develop microelectrode biosensor technology for real-time measurement of chemical signalling in the CNS. ND has invented a novel deposition technique based around silicate sol-gels to make highly sensitive biosensors for a much wider range of analytes including the purines plus acetylcholine and D-serine. The biosensor work has led to 3 patent applications, and a start-up company Sarissa Biomedical Ltd that markets these biosensors for researchers around the world.

ND is an acknowledged leader in the study of purine release, metabolism and actions. He has built up a comprehensive array of tools to study these important signalling agents. These tools (biosensors for ATP[ND5] and adenosine;[ND10] HPLC and optical methods for assaying E-NTPDase activity;[ND9,14,15] cell lines expressing specific E-NTPDases;[ND4] mammalian brain slice electrophysiology and measurements of synaptic transmission[ND9-11,14,15]) mean that he is ideally placed to provide a comprehensive characterization of the ability of the compounds generated within this study to block E-NTPDase activity in brain tissue, and to elucidate the extent to which these compounds affect other signalling mechanisms in the tissue.

ND has successfully supervised Masters Students, PhD students and postdocs during his career. Three students have completed their PhD under his supervision, a fourth student has submitted his PhD thesis and ND is currently supervising 2 PhD students who are in their 2nd year. The 4 year submission rate is 100% for these PhD projects. Furthermore, a MOAC Masters student will soon start his mini-project in ND's lab (co-supervised by Prof Richard Napier, Warwick HRI).

2. Background to research project

In addition to its key role in cellular metabolism the purine nucleotide ATP also acts as a potent extracellular messenger which is co-released under normal physiological conditions with a number of neurotransmitters including acetylcholine, norepinephrine, glutamate, γ-aminobutyric acid, and neuropeptide Y.[1] As a messenger ATP binds to P2 receptors which exist as two distinct families: the P2X ligand-gated ion channel (LGIC) family that is involved in fast excitatory neurotransmission and the metabotropic P2Y G-protein coupled receptor (GPCR) family. Once released, ATP is degraded to ADP, AMP, and adenosine by a family of ectonucleosidases (E-NTPDases), thus limiting the extracellular actions of the nucleotide(s) by enhancing its removal, as well as producing the pharmacologically active nucleoside, adenosine.

Numerous essential physiological processes are controlled by purinergic signalling.[2] Studies of purinergic signalling networks in tissues or cell cultures often use antagonists or agonist of P2 receptors and/or inhibitors of E-NTPDases in order to inhibit or enhance the activity of putative key players. However, current compounds bind with low affinity to these proteins and are very unselective which renders interpretation of biochemical results risky. New, highly selective inhibitors of E-NTPDases are desperately needed and would not only be useful tools for biochemical studies but would also have a huge therapeutic potential. To date, no atomic structures are available for any of the E-NTPDases.

3. Reseach project

Aims: The proposed programme of research aims to chemically synthesise ATP analogues, which are based on inosine and xanthosine nucleotides. These compounds will be tested against human E-NTPDase using *in vitro*, cell culture and tissue based assays in order to determine their potency and selectivity as inhibitors for individual E-NTPDase isoforms. E-NTPDase inhibitors with good selectivity and high affinities will be used to study purinergic signalling pathways in tissue and cell culture based systems.

***Originality*:** The proposed project uses the power of chemical synthesis to produce a highly diverse library of ATP analogues which are based on inosine and xanthosine nucleotides. Such compounds and derivatives have not been studied as E-NTPDase inhibitors to date and are therefore highly novel.

Feasibility & Adventure: The proposed chemistry is not highly risky and its convergent approach will produce a good selection of compounds. Prof Nicholas Dale is an internationally well established neuroscientist in the field of purinergic signalling. His expertise and the biological assays established in his research group will facilitate the biological study of the synthesised compounds. Data from these studies can immediately be fed back into the molecular design of the next generation of ATP analogues which then can be tested again. As a result, this iterative process is expected to furnish E-NTPDase inhibitors with high selectivity and potency. The challenge will be to produce compounds which are only selective for E-NTPDase and do not interact with other nucleotide and nucleoside recognising receptors in the cell, e.g. P2X, P2Y and P1 receptors. However, the proposed chemistry will produce a decent library of compounds with various profiles which will allow choosing the most successful ones. The biological study of the most successful compounds in cell culture and tissue based assays of purinergic signalling model systems, which include several regulatory and signalling components, will ultimately establish the usefulness of these compounds.

4. Scientific Strategy

Synthesis: Current E-NTPDase inhibitors can be grouped in two categories: 1) nucleotide analogues, more particularly P2 receptor agonist/antagonists and non-cleavable ATP derivatives and 2) the other molecules, most of which relate to the surinam-Evans blue dye families.[2] We propose to synthesise and biologically test hypoxanthine and xanthine nucleotide derivatives as novel E-NTPDase inhibitors. The structure of these oxidised nucleic acids combines the purine ring feature of adenine with a very similar (hypoxanthine) or identical (xanthine) H-bond donor/acceptor pattern of uracil in the 6-membered ring. Promising results with 5'-peptide derivatives of uridine as E-NTPDase inhibitors were recently presented.[3]

We are planning to use a systematic approach in the design of the xanthine and hypoxanthine derivatives whereby the parent nucleotides, *i.e.* inosine 5'-triphosphate (ITP, **1**) and xanthosine 5'-triphosphate (XTP, **2**), will be used as a starting point for the molecular design. We intend to successively chemically modify or to replace the nucleotide building blocks of the parent structures, *i.e.* the triphosphate chain, the ribose unit and the base, by structural surrogates in order to produce small compound libraries (Figure). The members of this small compound library will then be biologically tested and their pharmacological profile established. The most potent and selective inhibitors will be taken into the next modification round to produce a new series of compounds which then can be tested again.

Biology: Assays of activity against E-NTPDase activity will be conducted in a nested fashion. An optical assay of phosphate production in brain tissue will be utilized as a rapid screen to test their efficacy against native E-NTPDases[4] and identify "leads". Brain slices (hippocampus, cortex, cerebellum) will be placed in small chambers

and incubated with ATP either with or without the candidate compound. At varying time points samples will be taken and phosphate production measured and the effect of the compound on rates of phosphate production quantified. This rapid screen will form a first pass to identify compounds for more intensive investigation. Promising compounds will be further investigated by HPLC methods to determine their specific effects on ATP and ADP breakdown, giving information as to which of the E-NTPDases they may block in brain slices.[5] Cultured cell lines expressing cloned E-NTPDases (we shall concentrate on subtypes 1-3 as being of most relevance to neural function) will be used to quantify the IC_{50} of the compound for each of these enzymes (via either the optical assay method or HPLC). For those compounds that retain promise after this more in depth investigation, we shall use biosensors and electrophysiological methods to assay their actions in the context of physiological signalling. Utilizing hippocampal and cerebellar slices, we shall examine synaptic transmission (subject to purinergic modulation in both areas of the brain). ATP and adenosine biosensors will be utilized to examine how application of the compounds alter the dynamics of ATP breakdown and adenosine accumulation in real time[4,6] – both exogenously applied ATP, and endogenously released ATP. This will be simultaneously related to the real-time modulation of synaptic transmission by ATP and adenosine. Our electrophysiological methods will allow us to test whether the compounds alter other aspects of physiological function unrelated to their inhibition of the E-NTPDases – such as direct effects on transmitter release, neuronal excitability, ion channel gating, and activation or antagonism of purinergic receptors.

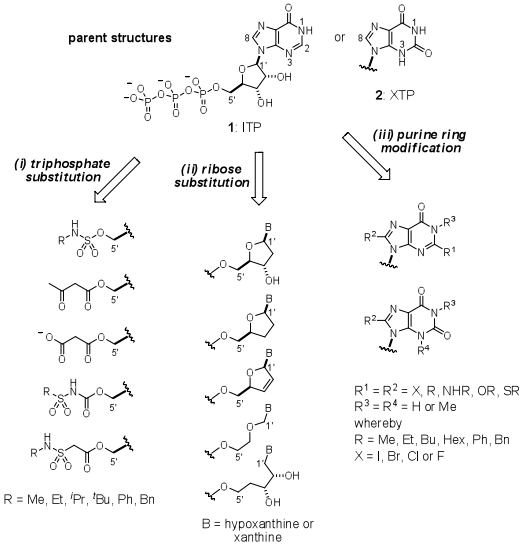


Figure. Structure-activity relationship study and target compound series.

5. Technical training

This project certainly involves a good deal of chemical synthesis and therefore is primarily aiming at postgraduates with this kind of interest. A potential student would be trained in modern methods of nucleotide synthesis which is certainly valuable considering the importance of this compound class. However, there is always feedback from the biological assays at early stages of the project and it becomes an iterative process rather than simply the synthesis of targets. Importantly, the student will conduct the biological assays him/herself and thereby acquire some useful biological skills as well.

6. Integration of the multi-disciplinary content

The supervisors have submitted a joint proposal to BBSRC last year which unfortunately was unsuccessful. Future application, e.g. to the Wellcome Trust, are planned.

The PhD student would spend the first phase of the project in the synthetic ML lab where he/she will get training in modern synthetic chemistry. Once the first compound is obtained he/she will then move to the biological ND lab and get training in all the required biological assays. Once the PhD student is more familiar with the chemical and biological techniques he/she will be able to take a more proactive role and divide up the time between chemistry and biology lab more individually depending on the progress of chemistry or biology. The project is of an iterative nature and so will be the interaction with the supervisors.

7. Justification of resources required

Due to the heavily experimental nature of the project a considerable consumables budget is required for the chemistry component, to purchase high purity solvents, reagents, silica gel for chromatography and inert gases. The biology consumables budget is substantial as well and covers the purchase of ATP sensors, high purity fine chemicals for HPLC, consumables for cell culture and electrophysiology. In total, £2000 pa is required.

8. Project management

ML and ND will take direct and overall responsibility for the research, reporting, supervision and progress of the proposed project. ML has a strong foundation in synthetic chemistry and will advise and practically support the PhD student on the design and synthesis of the target compounds. Given the huge expertise in purinergic signalling research ND will design and supervise the biological assays that the synthesised compounds will be subjected to. ML, ND and the PhD student would have a project review meeting each month.

9. Plan for what the PhD student will do in the first 3-6 months

- month 1: Familiarisation with project and chemical laboratory and survey of literature.
- months 2-4: Synthesis of first 1-3 compounds in which the triphosphate chain has been replaced (Figure).
- months 5-6: Familiarisation with biological laboratory and assays; biological study of synthesised compounds.

10. References

ML1 M. Lochner, A. Geneste, M. Hesse, *Helv. Chim. Acta* **1998**, *81*, 2270.
ML2 Y. Li, K. Popaj, M. Lochner, H. Geneste, R. Budriesi, A. Chiarini, C. Melchiorre, M. Hesse*, Il Farmaco* **2001**, *56*, 127.
ML3 M. S. Robillard, B. A. J. Jansen, M. Lochner, H. Geneste, Y. Li, J. Brouwer, M. Hesse, J. Reedijk, *Helv. Chim. Acta* **2001**, *84*, 3023.
ML4 M. Lochner, L. Mu, W.-D. Woggon, *Adv. Synth. Catal*. **2003**, *345*, 743.
ML5 M. Lochner, M. Meuwly, W.-D. Woggon, *Chem. Commun*. **2003**, 1330.
ML6 A. J. Thompson, M. Lochner, S. C. R. Lummis, *Br. J. Pharmacol*. **2007**, *151*, 666.
ML7 M. Lochner, A. J. Thompson, S. C. R. Lummis, *Biophys. J.*, accepted for publication.
ND1 N. Dale, *J. Neurosci*. **2005**, *22*, 10461.
ND2 N. Dale, *J. Physiol*. **1998**, *511*, 265.
ND3 N. Dale, D. Gilday, *Nature* **1996**, *383*, 259.
ND4 K. Massé, R. Eason, S. Bhamra, N. Dale, E. Jones, *Genomics* **2006** *87*, 366.
ND5 E. Llaudet, S. Hatz, M. Droniou, N. Dale, N. *Anal. Chem*. **2005**, *77*, 3267.
ND6 N. Dale, A. V. Gourine, E. Llaudet, D. Bulmer, T. Thomas, K. M. Spyer, *J. Physiol*. **2002**, *544*, 149.
ND7 A. V. Gourine, E. Llaudet, N. Dale, K. M. Spyer, K.M. *Nature* **2005**, *436*, 108.
ND8 A. V. Gourine, E. Llaudet, N. Dale, K. M. Spyer, *J. Neurosci*. **2005**, *25*, 1211.
ND9 B. G. Frenguelli, G. Wigmore, E. Llaudet, N. Dale, *J. Neurochem*. **2007**, *101*, 1400.
ND10 B. G. Frenguelli, E. Llaudet, N. Dale, *J. Neurochem*. **2003**, *86*, 1506.
ND11 T. Pearson, F. Nuritova, D. Caldwell, N. Dale, B. G. Frenguelli, *J. Neurosci*. **2001**, *21*, 2298.
ND12 N. Dale, T. Pearson, B. G. Frenguelli, *J. Physiol*. **2000**, *526*, 143.
ND13 R. A. Pearson, N. Dale, E. Llaudet, P. Mobbs, P. *Neuron* **2005**, *46*, 731.
ND14 M. Wall, A. Atterbury, N. Dale, *J. Physiol*. **2007,** *582*, 137.
ND15 M. Wall, N. Dale, *J. Physiol.* **2007**, *581*, 553.
1 H. Zimmermann, *Pflugers Arch*. **2006**, *452*, 573.
2 H. Ullmann, S. Meis, D. Hongwiset, C. Marzian, M. Wiese, P. Nickel, D. Communi, J.-M. Boeynaems, C. Wolf, R. Hausmann, G. Schmalzing, M. U. Kassack, *J. Med Chem*. **2005**, *48*, 7040.
3 A. Brunschweiger, J. Iqbal, A. Scheiff, M. N. Munkonda, J. Sevigny, A. F. Knowles, C. E. Müller, *conference poster*, 8th international symposium on adenosine and adenine nucleotides, Ferrera, Italy, May 2006.
4 B. G. Frenguelli, E. Llaudet, G. J. Wigmore, N. Dale, *J. Neurochem.* **2007,** *101*, 1400.
5 M. J. Wall, N. Dale, *J. Physiol.* **2007,** *581,* 553.
6 M. J. Wall, A. Atterbury, N. Dale, *J. Physiol.* **2007,** *582*, 137.

MOAC project 12

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: Microfluidic devices for probing DNA-protein interactions in cells: A new interface across Engineering, Systems Biology and DNA Chemistry.

Supervisor name	Department/email/phone number	Supervisor's advisor*
Dr James Covington	Microsystems and Power Devices Group School of Engineering, j.a.covington@warwick.ac.uk x74494	
Prof Georgy Koentges	WSB , g.koentges@warwick.ac.uk , 07837 207066	

Project proposal:

1.Track record of supervisors in research and supervision

Prof Georgy Koentges, appointed in 2007 as experimental co-director of WSB is an international expert in single cell transcriptome profiling and dynamic measurements of gene activity in the area of vertebrate developmental biology and evolution. After studies in comparative anatomy, evolution and philosophy in Tuebingen and Freiburg, he was trained as a molecular embryologist by Prof A Lumsden, FRS, was awarded the T.H.Huxley medal for the best UK PhD thesis in the life sciences, then worked with Prof Dulac and Nobel-prize winner Richard Axel on deciphering the molecular anatomy of the olfactory system through gene-targeting approaches, published in *Cell*, *Science* and *Neuron*. Since starting his own lab in 2001 (appointed as Senior Lecturer at UCL) he published in *Science*, *Nature* etc, developed novel strategies to expression profile single cells from tissues and has been addressing some fundamental questions pertaining to the molecular and cellular evolution of shapes (recently outlined in Koentges,*Nature*, Feb 8th, 2008). As part of this goal he has more recently turned his attention to understanding the logic of *cis*-regulatory coding in the genome as measured in single cells and has developed with Keith Vance, Sascha Ott and others at WSB an imaging pipeline for single cell tracking of gene activity patterns. His work is conducted with a number of theory collaborators on campus and in the US and is funded over the next four years by two Programme grants of the Wellcome Trust as well as a Human Frontiers science programme grant with collaborators at Stanford and Israel (ranked 6/780 international grant applications). He has supervised 4 PhD students (full-time and part-time) resulting in publications in *Nature* and other journals. All PhDs were completed in the time allocated/available through studentship (full-time students)or grant funding (for part-time students). Current work of a (part-time) PhD student is being written up for *Nature* and is the biological motivation of the currently proposed more technical application across Engineering and WSB.

Dr James Covington is an Associate Professor in the School of Engineering. He was appointed as a lecturer in 2002, where he joined the Electrical and Electronic Division within engineering. He is a specialist in micro-systems, bio-chemical sensors and semiconductor devices. He has considerable experience in the field of micromachining, silicon fabrication, chemical sensor design and ASICs. For the last 10 years Dr Covington has been developing chemical and biological sensors for detecting a variety of environmental pollutants and biological agents, applying a wide range of electronic and MEMS techniques in the development of these novel sensors. More recently, Dr Covington has become interested in creating devices for highly novel medical and BioMEMS applications and is developing, for example, innovative micro-systems for the blood sampling and analysis. Dr Covington is in overall control of the micro-fabrication and MEMS facilities (including the School of Engineering clean room), which is supported by two technicians and a number of PhD students performing a broad remit of research activities. Presently Dr Covington has 6 active funded research projects totalling over £1M. In addition, he has previously benefited from substantial equipment investment through successful SRIF and Royal Society Wolfson awards. He presently supervises 7 Engineering PhD students. Engineering PhD students typically submit between 3-5 journal papers (in world leading publications) and up to 5 conference papers. He is eager to join forces with the molecular biologists at WSB on a truly interdisciplinary project under MOAC and bring the capabilities of microfluidic engineering into MOAC's portfolio.

2.Background to research project and timeliness

Rendering the dynamics of biological behaviour measurable at the scale at which it happens, i.e. in single cells in real-time is a key objective of 21st century biology in general and systems biology in particular. This requires the establishment of measurement and manipulation platforms for single cells (recently outlined in *Koentges, Nature Feb 8, 2008*). Such platforms have to have the capability of multiplexing, i.e. many parameters need to be measured in hundreds and thousands of cells; cells that have to be manipulated with complex libraries of thousands of (DNA as well as chemical) compounds. This is needed in order to probe and traverse the combinatorial complexity of regulatory phenomena in living beings and is a key objective of current, genome-based basic research as well as translational medicine. Koentges' team at **WSB** has established an image analysis pipeline (with the help of Drs. Ott, Bretschneider, Vance and others) that is based on the Cellomics KSR machine which allows us to make microscopic images of thousands of cells over several days. **We do not have the capability yet to individually address these cells and manipulate their measured dynamics in a massively parallel fashion.** This requires the ability to seed and trap cells into thousands of micrometer-sized containers that are **transparent** for UV light and fluorescence and that are **sufficiently hospitable** for cells to grow in over several days. **James Covington, who heads the Microsystems Group at Engineering** has significant expertise and capabilities in the area of microfabrication that will allow us to jointly overcome current limitations in the field. The types of experiments that can become possible through such multiplexing device made through microfabrication are expected to lead to a step-change in our ability to manipulate biological systems. Familiarity with all the tools of the microfabrication engineer as well as the molecular biologist are going to be significant transferable scientific skills both in the sciences as well as biomedical industry.

3.Research project

Aims

The present PhD project that cross-cuts department/discipline boundaries attempts to establish a very **first prototypic microwell transfection device** to be placed into our existing Cellomics imaging platform that will allow us to render thousands of cells individually addressable through DNA or chemicals and imageable for effects of such treatments. Such proposal is only possible through developing a close interface between three disciplines, Biological Sciences, Systems Biology and Engineering. It takes advantage of existing microfabrication facilities in engineering as well as the established image analysis pipeline for single cells and protocols for the transfection and culture of stem cells established in the Koentges lab at WSB that is internationally unique. The present proposal will allow us to establish such link for the first time on campus in the areas of microengineering and systems biology - that we believe is critical for our future success in this area internationally and locally. Novel devices will be equipped **and tested with complex libraries of DNA constructs** that will be purchased or custom-generated by international companies and US colleagues (Prof. Gao, U o H). Beyond the design and production of a device prototype through several manufacturing and experimental testing iterations for which we will use libraries of DNA constructs. The student will learn how to **manufacture microwell devices** and to **adapt them to biological applications in an interative manner**, to **make single-cell measurements** in these devices and to **generate (classes of) DNA constructs** to load into these devices, based on a rational strategy to change binding strengths of trasncription factor binding sites on regulatory regions through molecular biology.

4. Research plan: Methodology and two phases.

We envisage this pilot project to consist of two major phases, **Phase 1.** First design and testing of a massively parallel transfection device by using already available DNA construct libraries, this is expected to cover the first (Month 1-18) (JC,GK). **Phase 2:** After that his focus will be on generating DNA construct libraries to be placed into the device, tested and the resulting fluorescence parameters being measured inside the Cellomics KSR (GK, XG).

Phase 1.Microwell transfection and growth device – design, manufacture and testing

Our intention is not to create a highly sophisticated fluidic system that has individual channels to each well, but to create a device that can be "loaded" with DNA in parallel through a staged process. We propose to first fabricate a fluidic package that will hold the cells and be used to monitor any fluorescence. This will be made by depositing and patterning a poly-lysine+ cell adhesion layer on the surface of a glass plate. This glass will then be loaded with cells and washed, thus only cells on the adhesion layers will remain. Then a plastic package will be aligned and fitted over the glass plate. Clearly, care will be required to ensure that the bonding process, where the glass is adhered to the plastic, does not affect the cells. This plastic package will contain a number of chambers, where the cells will reside. In addition, there will be a number of needle like structure on the other side, thus at this stage the sample is open to the environment. Then we propose to fabricate a second plastic micro-package containing a simple set of chambers. These will be loaded in the first instance with different DNA samples available in the Koentges lab. Necessary techniques are common place within the biological sciences and all necessary stem cell transfection and cellular growth protocols are firmly established in the Koentges lab. Over the top of this loaded chamber will be fitted a thin plastic sheet (somewhat like Clingfilm). It is intend that this layer is not, as such deposited, but more "shrink rapped" over the surface. Finally

the two packages will be brought together, where the needle point pierce the cling film layer and the DNA sample flows into the cell loaded chamber. This concept is shown below in figure 1. Here figure 1a shows the glass-plastic fluidic concept and figure 1b shows how DNA will be transferred to the cell loaded chamber for transfection of the cells housed in these chambers.

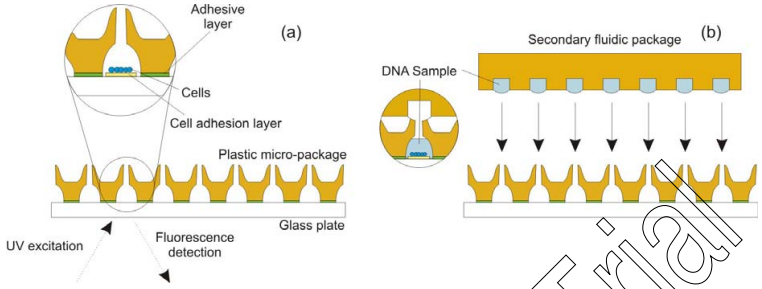


Figure 1: (a) Glass-plastic fluidic package and (b) DNA transfer

The fluidic chambers will be fabricated with a dimension of typical 200 – 500 μm. We propose that each package will contain between 1000 and 10,000 different chambers. Though it is unlikely that we will have 100% yield, by using multiple samples of the same type we should ensure good results. In addition, we have methods the coat the plastic to ensure bio-compatibility and survive most sterilisation processes. The plastic component of the system will be constructed using 3D rapid micro-manufacturing techniques, based within the School of Engineering. Fabricating structures and the scales proposed here is achievable (minimum features sizes are well below these requirements) using a microfabrication machine that allows rapid turnover between design and manufacturing of microwell devices within a day. In direct 3D rapid micro-manufacturing, a 3D CAD (Computer Aided Design) model of the object is first created. This CAD model is then sliced horizontally into a series of 2D images that represent multiple cross-sections of the 3D object. These layers are translated into appropriate control and positioning co-ordinates and are cured layer by layer (cross section by cross section), into a photo-curable resin. After each layer has been cured into the resin, the object is moved vertically to allow an uncured layer of resin to cover the previously cured layer. This is repeated until the object is formed.

Phase 2: After that his focus will be on generating DNA construct libraries to be placed into the device, tested and the resulting fluorescence parameters being measured inside the Cellomics KSR (GK, XG).

To broaden the scope of the DNA libraries to be spotted into the transfection device, the student will go to UoH, the laboratory of our collaborator Prof Gao and learn the *ab initio* assembly of DNA and the massively parallel synthesis microfluidics (Tian et al., Nature 2004) and will synthesize the constructs that he plans. It is envisaged he will make detailed construction plans over several (3-4) months at WSB, the goes to UoH for 3 months and will spend the remainder of his PhD iteratively testing the constructed DNAs carrying various modificaitions of transcription factor binding sites. It is important to point out that the way how to do this, what exactly to do etc will be closely guided by Prof Koentges and his staff and postdocs. In our view, the generation of microwell devices, their testing and application within a bigger WSB cell tracking project are meritable research aims on their own, given the non-trivial nature of the procedures involved. Detailed biological explanations about the particular project can be provided or can be gleaned from the course work of the CH927 module.

Feasibility

Using this technique it is possible to design and fabricate a package in less than a day, with a repeat package fabricate in a number of hours. Thus it is a simple process to develop and test new concepts and ideas. Furthermore, there is extensive training available on both CAD and the use of the machines. **A student with keen interest in experimentation and manufacturing but without previous experience in this field can learn within a few weeks the operations, software etc required for running these machines successfully. The present proposal is solely based on current capability.** This massively parallel transfection and imaging apparatus is then placed into the Cellomics KSR machine, which is temperature and CO_2/O_2 controlled and viability will be tested over the usual assay period of 3 days and many cycles are envisaged.

6.Technical training – see above and below.

7. Integration of the multi-disciplinary content (this may include a track record of collaboration between supervisors and an indication of how the student and supervisors will interact).

The present project aims to complement the existing (and working) imaging pipeline we have built at Warwick with an equally powerful manipulation platform that allows single cells to be trapped into small containers in order to transfect these with complex libraries of DNA constructs. **This is part of an ongoing collaboration between James Covington and Georgy Koentges that brings together state-of-the art local microfluidic fabrication expertise in Engineering with the molecular biology manipulation at WSB**. The current interdisciplinary PhD project aims at developing a transfection machine that will enable us to conduct massively parallel (but independent) DNA transfection experiments of single cells. Such device will be integrated with our existing imaging platform that captures the functional dynamic readouts of cis-regulatory module combinations. The student will fabricate the device prototypes under the guidance of Covington, will then test them in the Koentges lab on the Cellomics KSR for their optical properties, cell survivability etc which will lead to numerous iterations and improvements over the whole period of the PhD. The student will also become involved in the construction of the DNA that is spotted into this device, using novel microfluidic chemistry under the guidance of Prof Gao. This project will grant the student the chance to also work with our close collaborator **Prof Gao**, University of Houston, Texas, a pioneer in the microfluidic mediated assembly of DNA constructs (Tian et al. Nature 2004), a technology platform we will be working with in the future and that the student can immediately tap into by spending time at Prof Gao's lab. Success of this experimental PhD is not contingent upon this collaboration but allows the student to get access to a wider network of investigators of international standing. In the Koentges lab the student will learn how to design DNA constructs that contain regulatory regions, will learn how to modify the transcription factor binding sites contained in these regions in a systematic manner and assay the resulting phenotypes in vivo. For this the student will learn advanced cloning, molecular biology and tissue culture techniques. At WSB his work will be part of an attempt to understand and model the underlying logic functions of gene-regulatory regions with state-of-the-art statistical mechanics approaches, gleaned from physics, statistics and other disciplines, previously outlined in the CH927 module.

8.Justification of resources required

We request the full amount of financial support available to the student to be shared 50/50 among the two supervisors. This is to cover some minor lab expenses in microfabrication, molecular biology and DNA synthesis as well as travel/accommodation for the student to Prof Gao's lab at UofH, Texas. Prof Koentges will commit the remainder of the actual consumables costs on his side towards this project from available grants (about 10K/yr) and Dr Covington will do the same on the engineering side. Please note that this is a significant financial commitment on our side.

9. Project management

We expect this project to be undertaken 60% at Warwick Engineering and 40% at Warwick Systems Biology lab at Gibbet Hill. Supervisors meet the student each week, the student takes part in all regular group meetings and activities to learn techniques and the different ways of thinking across engineering and molecular biology. The possible component of work with Gao will fall under the WSB tranche.

10. Plan for what the student will do in the first 3–6 months,

See above. Student will familiarize himself with microfabrication technology over several weeks, then generate the first sets of prototypes, go back to the WSB lab and test them. He will learn how to use the Cellomics KSR imaging set-up at WSB. Within the first 6 months he will have created a number of workable prototypes. Within 16 months he will have generated DNA construct libraries with Prof Gao and have done the first sets of large-scale measurements.

MOAC project 13

MOAC PhD-project proposal

Submit up to FOUR pages including this header page to Alison Rodger (a.rodger@warwick.ac.uk) by Sunday 1st June, 2008).

Project title: Membrane Proteins in their Natural Environment: Solid-state Nuclear Magnetic Resonance and Molecular Dynamics Simulation for Detwermination of Structure and Dowynamics

Supervisor name	Department/email/phone number	Supervisor's advisor*
Dr. Ann Dixon	Chemistry/ann.dixon@warwick.ac.uk/50037	Prof. Alison Rodger
Dr. Steven Brown	Physics/s.p.brown@warwick.ac.uk/74359	
Prof. Mike Allen	Physics/m.p.allen@warwick.ac.uk/74415	

Project proposal.

(1) Track record of supervisors in research and supervision.

***Dr. Ann Dixon* [AMD]** (Assistant Professor, Chemistry, MOAC lecturer in Chemistry, University of Warwick since 2005). Extensive experience in the investigation of the structure, folding, and interactions of integral membrane and membrane-associated proteins. Current research is focused on investigation of transmembrane (TM) helix-helix interactions and their roles in membrane protein structure, biological function and disease. 15 primary literature papers and one book chapter. Research funding of £635K (from CRUK, DUK, and MRC) obtained while at Warwick as PI. Member of Midlands Biophysics Network and Midlands Membrane Protein Network, and secretary of the RSC Biophysical Chemistry Group. Currently supervising two postdoctoral fellows, two Chemistry PhD students (both in second year of PhD), and two MOAC PhD students (jointly supervised with Dr. Graham Ladds and Prof. Colin Robinson). First MOAC PhD student due to complete in Summer/Autumn 2008. To date, 2 publications with PhD students (2 more submitted/in preparation).

***Dr. Steven P. Brown* [SPB]** (Associate Professor, Physics, University of Warwick since 2002). International reputation for developing new solid-state NMR techniques and their application, e.g. for studying hydrogen-bonding interactions in pharmaceuticals. 50 primary literature papers (16 in JACS/Angew. Chem.) (h-index = 22). Research funding (EPSRC, Leverhulme, Royal Society) at Warwick as PI of £5M: chair of the National Management Committee of the £3.7M EPSRC/BBSRC funded 850 MHz UK solid-state NMR facility, to be sited at Warwick, anticipated opening Summer 2009. His first PhD student completed in January 2008 (within 3.5 years) and is now carrying on solid-state NMR research as a PDRA in St Andrews. Currently supervising 3 PhD students (supported in part by AstraZeneca and GSK). To date, 7 publications with Warwick PhD students. Lecturer in MOAC.

***Prof. Michael P Allen* [MPA]** (Professor, Physics, University of Warwick since 2001). Previously at Bristol University (Lecturer 1985, Reader 1992, Professor 1997). 30 years experience in molecular simulation, over 100 primary literature papers, one book, two edited collections. Founding director of Centre for Scientific Computing at Warwick; founding organiser and principal lecturer (with DJ Tildesley and JHR Clarke) of the CCP5 annual graduate summer school "Methods in Molecular Simulation" (1994-2001). Supervised 17 PhD students to date, of whom 11 have completed and 3 are currently in progress. 4 PhD students have gone on to academic careers, 1 has a postdoctoral position, and MOAC PhD student John Grime (jointly supervised with Pat Unwin) will take up a post-doctoral position at Uppsala following submission of his thesis later this month. 34 papers published with PhD students plus 2 with final-year project students.

(2) Background to research project.

A key element of a cell is its lipid membrane; as well as sequestering the contents of the cell, the membrane, and in particular membrane proteins that are embedded in the membrane, vitally regulate how the cell interacts with its environment. There is, hence, great interest in medicine and industry to obtain structural information about these membrane proteins in order to understand disease pathways and target new pharmaceuticals. However, while membrane proteins account for 30% of proteins encoded by the human genome, only approximately 100 membrane protein structures have been successfully solved (as compared with the over 30,000 structures for soluble proteins). Indeed, discoveries related to membrane protein structures have led to three Nobel prizes since 1987. This vast discrepancy in the number of known structures is due to experimental difficulties, as membrane proteins are difficult to crystallise for X-ray studies, and can produce large aggregates that exceed the size limit for solution-state nuclear magnetic resonance (NMR) spectroscopy.

In contrast, solid-state NMR is uniquely placed to study membrane proteins in their natural environment, a lipid bilayer. Solid-state NMR is increasingly used in structural studies of biosolids, as evidenced by a number of recent publications in leading journals such as PNAS and Nature [1-6]. In particular, solid-state NMR has much potential as an atom-specific probe of structure and dynamics for key proteins, including membrane proteins. Thus far, this technique has mainly been used to study membrane proteins containing isolated pairs of labelled nuclei, however recent technique developments are aimed at obtaining many distance constraints for multiply labelled samples. Another method that is proving invaluable in the analysis of membrane proteins is computational modelling, especially when resolving uncertainties in the global protein fold (e.g. α-helix packing) and oligomer structures. The simple two-stage model of membrane protein folding (insertion of individual helices into the membrane followed by association into a tertiary structure) [7] has led to the development of computational tools of varying sophistication such as energy minimization and molecular dynamics (MD). These may take account of packing considerations, van der Waals interactions, electrostatics and hydrogen bonding. One of the simplest approaches is a global conformational searching algorithm (called CHI - see * below) followed by MD in vacuo [8, 9]. More recently, MD in a full lipid bilayer environment has been used to test proposed structures, discriminating between different helix packings [10, 11], predicting helix dimer structures [12], and studying the effects of hydrophobic mismatch [13]. A well-established protocol is to insert a protein in a pre-formed bilayer such as DMPC [14] and carry out a coarse-grained self-assembly simulation [15] prior to full MD with a standard package such as GROMACS or CHARMM. Various techniques can be used to speed up the process [16].

The transmembrane domain of Glycophorin A (GpA; Fig 1a) is a simple and well-characterised example of the type of protein that we wish to investigate in this project: it contains a single transmembrane domain (Fig.1a) that inserts into membrane bilayers as a stable α-helix and drives protein dimerisation (Fig.1b) [17]. The dimer structure of GpA was first predicted by the program CHI, mentioned above, [8, 18] and later confirmed by solution-state NMR experiments in micelles [19]. More recently, structural data in membrane bilayers has also been collected using conventional solid-state NMR methods, providing further insight into the amino acids that stabilise the GpA dimer [19-21]. For example, a sequence motif that is critical for correct folding of GpA is the GG4 motif, where glycine residues located every four amino acids localise to one face of the helix and pack against the other helix (Fig.1b) [22].

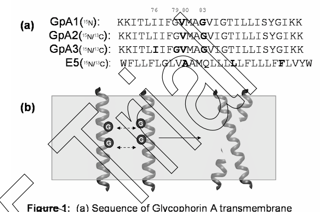


Figure 1: (a) Sequence of Glycophorin A transmembrane domain peptide, highlighting the sites of isotopic labels, and the E5 transmembrane domain peptide. (b) GpA α-helices, containing the GG4 motif, interact to form a dimer in the membrane bilayer.

***Aim*:** The aim of this project is to combine complementary techniques in biological sample preparation, novel solid-state NMR methodology, and atomistic to coarse-grained molecular modelling to extract structural information for membrane proteins in a biologically relevant environment. NMR data will provide atomic-level information which can be fed into the CHI program to produce 3D structural models. Coarse-grained modelling can then provide a bigger picture of the folding and assembly process.

***Preliminary data*:** In initial work done in a mini-project last year, experimental protocols were established in the Dixon and Brown groups for the preparation of membrane proteins (prepared by chemical synthesis with specific isotopically labelled sites) within hydrated DMPC vesicles, suitable for analysis by solid-state NMR. Initial data was collected on GpA, which acts as an ideal test system with which to assess our protocols. Specifically, we synthesised a 30-residue peptide corresponding to the transmembrane domain of GpA containing ^{15}N and ^{13}C isotopic-labels on Val_{80} and Gly_{83} (Fig.1a). The peptide was incorporated into fully-hydrated DMPC bilayers, and packed into a solid-state NMR magic-angle-spinning (MAS) rotor. A ^{13}C spectrum of the sample is shown in the upper half of Figure 2, where peaks from both the peptide and the lipid can be seen. In order to remove any peaks due to the lipid, thus simplifying the spectrum, we then employed double-quantum filtered (DQF) solid-state NMR techniques. This approach yielded the spectrum shown in the lower half of Figure 2, where only peaks due to the labelled amino-acid residues of the transmembrane protein are seen, and the natural abundant lipid signals are filtered out. The assignment of the carbon atoms in the two labelled amino acids (Fig. 2, right panel) are shown next to each peak in the spectrum (Fig. 2, lower spectrum).

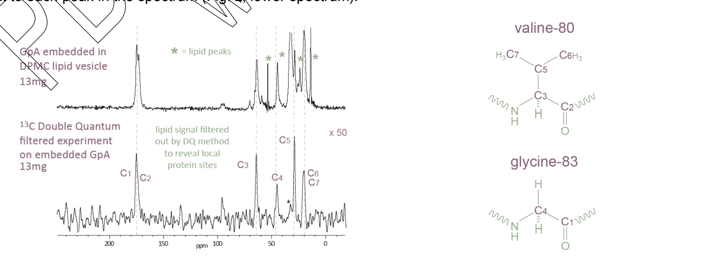


Figure 2: *Upper spectrum*: ^{13}C NMR spectra of the GpA/DMPC sample. *Lower spectrum*: ^{13}C DQF spectrum containing only peptide peaks. Assignments are listed next to each peak. *Right panel*: Two ^{13}C labelled amino acids, and corresponding carbon assignment.

---

*: CHI="**C**NS searching of **H**elix **I**nteractions", where CNS="**C**rystallography and **N**MR **S**ystem"

Our preliminary data speaks to the feasibility of this project for a MOAC PhD student. We can now show that we can synthesize, purify and incorporate protein transmembrane domains into synthetic lipid bilayers. Furthermore, we can pack our membrane-embedded peptide samples into MAS NMR rotors and acquire meaningful data on these samples. The NMR peaks are sufficiently narrow to minimize spectral overlap, while being representative of a solid sample.

Proposed work in PhD: We would now like to push this technology forward. As mentioned above, solid-state NMR methods for determining dipolar couplings and hence distances are well established for samples containing isolated pairs of labelled nuclei, but new advanced methods for obtaining many distance constraints for multiply labelled samples are being developed at Warwick and elsewhere. Such methods will reduce the amount of time it takes to obtain data as well as reducing the cost. Therefore, a key part of the proposed PhD research project is to evaluate the advanced NMR methods using further GpA peptides with progressively greater degrees of ^{13}C and ^{15}N labelling (see GpA 3, Fig.1a). Once reliable methods are established for the GpA standard, work can begin on proteins of high biological importance. At this point, there is scope for the student to select a protein of specific interest. As a suggestion, we have selected the transmembrane domain of the E5 protein of Bovine Papillomavirus (Fig. 1a), a protein that is of great interest in cancer and is currently under investigation in the Dixon group. Structural information for this peptide would provide insight into viral cancers and signalling, as well as the field of therapeutic protein design. This peptide would be synthesised, initially labelling three positions along the chain thought to be important for its function.

To guide and inform this work, the student will also use a variety of molecular modelling techniques, ranging from atomistic modelling of protein structure up to coarse grain modelling of folding/oligomerisation events. A combination of techniques will be used in both an exploratory way (broadening the range of simulated configurations by replica exchange / parallel tempering, and comparing the simulation outputs with NMR data) and an experiment-driven fashion (narrowing the range of simulated configurations by incorporating constraints from the NMR in the simulations themselves).

(3) Research project

Originality: The field of membrane protein structural biology still lags far behind that of soluble proteins, and therefore much more work is needed. The use of solid-state NMR to tackle this problem is not new, however to date only partial structural data has been provided. Our approach in this project is to combine our expertise in membrane protein biochemistry, solid-state NMR, and atomistic to coarse-grained computational modelling to produce a well-rounded picture of proteins of high biological importance.

Feasibility: Our preliminary data (above) speaks to the feasibility of this project, and we now have suitable protocols for some of the more difficult aspects of the work such as sample preparation, reconstitution into lipids, and loading the MAS rotor. We have identified the conditions (temperature, concentration) which yield narrow peaks in the solid-phase spectrum, and have successfully employed DQF experiments for preliminary analyses.

Adventure: Membrane protein structure determination is widely seen to be the next frontier of structural biology, and is still viewed to be a fairly high-risk endeavor due to difficulties in working with these hydrophobic proteins. We have worked together over the past year to produce protocols and preliminary data that removes some of this risk, however any work in the field of membrane protein analysis can be called "adventurous".

(4) Scientific strategy

This is an interdisciplinary project that brings together supervisors from Chemistry and Physics with internationally recognised expertise in biophysical chemistry, solid-state NMR and modelling, in an integrated project to gain new insight into membrane protein structure and dynamics. The project, thus, fits closely to the identified focus of MOAC2 with membrane proteins being one of the four named biomolecular assemblies whose structure and function is to be targeted. Moreover, the project explicitly brings together model building with experiments that are at the cutting edge of experimental and instrumental design. The project supervisors are integral members of interdisciplinary centres that have received considerable recent University investment, namely the new Milburn House Magnetic Resonance centre (SPB & AMD) and the Centre of Scientific Computing (MPA, with SPB's group also being current users of the computational resources).

(5) Technical training.

In the Dixon group, the student will obtain training and experience in protein biochemistry and analytical techniques such as high performance liquid chromatography [HPLC], mass spectrometry [MS], electrophoresis [SDS-PAGE], circular dichroism [CD], and fourier-transform infrared spectroscopy [FTIR]. Use of common instruments such as UV-Vis spectrophotometers and centrifuges will also be obtained. In the Brown group, the student will obtain NMR experience at one of the largest solid-state NMR facilities in the world. Experience will include running of standard experiments, as well as methods development and pulse programming. In the Allen group, the student will become proficient in a wide variety of molecular modelling techniques and will learn to use the results from modelling to guide and inform his/her work. During the PhD, the student will gain training and expertise of presenting results orally and in written form, and will gain on-the-job experience through taking an increasing role in driving forward an interdisciplinary project. In addition, as is common across MOAC, the student will follow the Warwick Postgraduate Certificate in Transferable Skills in Science.

(6) Integration of the multi-disciplinary content (this may include a track record of collaboration between supervisors and an indication of how the student and supervisors will interact)

This project builds on preliminary results from a joint project between SPB & AMD that began in 2007 with the supervision of a Systems Biology mini-project student and a £9.5k award from the Warwick Research and Development Fund. In addition to interacting directly with the named supervisors and their groups (see Section 8), the PhD student can expect further support from an EPSRC LSI post-doctoral fellow, Dr. Johanna Becker. Dr. Becker has a great deal of experience in membrane protein structure determination by solid-state NMR, and will start her 3-year fellowship in the solid-state NMR group in Summer 2008.

(7) Justification of resources required.

For the synthesis and purification of membrane-embedded protein samples, a contribution of £1000 pa is requested for chemicals and consumables. A contribution to NMR running costs (cryogens, replacement rotors, stators, etc) of £500 pa is also requested. For the computational part of the project, a sum of £690 pa is requested. This comprises the standard charge of £150pa for users of the Centre for Scientific Computing desktop, plus an allocation of £350pa for time on the CSC high-performance cluster (roughly giving 8 nodes for about 42 days running per year).

(8) Project management.

The overall split of time between the three groups can be quite flexible and may evolve over time; however we initially expect that the split will be as follows: 50% AMD; 25% SPB; 25% MPA. The student will prepare the samples and carry out solid-state NMR experiments under the supervision of AMD and SPB, as well as experienced members of both groups [Dr. Andrew Beevers (PDRA, AMD group) and Amy Webber (PhD, SPB group)], respectively. The modeling studies will be done under the direct supervision of MPA. Progress will be closely monitored through a number of meetings with various groups. The MOAC PhD student will meet once a week with the AMD and SPB group, and one-on-one with all supervisors at least every two weeks (depending on the work program). The student will meet with all three supervisors once per month, and with their advisory committee according to standard MOAC regulations.

(9) Plan for what the student will do in the first 3–6 months (12-24 weeks).

The plan below is fairly rigid in the first 12 weeks, but then must be led by results from that point forward.

Week	Experimental	Theoretical
1-2	Literature survey, place order for synthesis of quadruply-labelled GpA peptide.	Literature survey
3-5	Training on HPLC followed by purification of GpA peptide.	
5-6	Training on mass spectrometers and SDS-PAGE, followed by characterisation of pure peptide by these methods.	
7-8	Reconstitution of peptide into lipid vesicles (bilayers). Training on CD and FTIR followed by analysis of peptide-lipid sample by both methods.	
9-10	Load sample into MAS rotor and acquire first NMR spectra.	Training in CHI followed by GpA dimer search using CHI.
11-12	Further NMR training and analyses of quadruply-labelled GpA sample.	Analysis of CHI data. Begin energy optimisation and MD runs at various temperatures starting from CHI-generated configurations. Continue with optimization and further MD runs.
13	Evaluate NMR data: if successful, place order for next peptide (e.g. multiply labelled E5 peptide or other): should progress according to approximate timetime for GpA above; if more work needed, continue with NMR analysis/methods development.	

(10) References

1. Egawa, A., et al., Proc. Natl. Acad. Sci. USA, 2007. **104** : p. 790-795.
2. Ferguson, N., et al., Proc. Natl. Acad. Sci. USA, 2006. **103**: p. 16248-16253.
3. Franks, W., et al., Proc. Natl. Acad. Sci. USA, 2008. **105**: p. 4621-4626.
4. Iwata, K., et al., Proc. Natl. Acad. Sci. USA 2006. **103**: p. 18119-18124.
5. Lange, A., et al., Nature, 2006. **440**: p. 959-962.
6. Williamson, P., et al., Proc. Natl. Acad. Sci. USA, 2007. **104**: p. 18031-18036.
7. Popot, J.L. and D.M. Engelman, Biochemistry, 1990. **29**(17): p. 4031-4037.
8. Adams, P.D., D.M. Engelman, and A.T. Brunger, Proteins, 1996. **26**(3): p. 257-261.
9. Brunger, A.T., et al., Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-921.
10. Barth, P., J. Schonbrun, and D. Baker, Proc Natl Acad Sci USA, 2007. **104**: p. 15682.
11. Ivetac, A. and M. Sansom, Eur Biophys J, 2008. **37**: p. 403.
12. Soumana, O., N. Garnier, and M. Genest, Eur Biophys J, 2007. **36**: p. 1071.
13. Petrache, H., et al., J. Mol. Biol., 2000. **302**: p. 727-746.
14. Ivetac, A., J. Campbell, and M. Sansom, Biochemistry, 2007. **46**: p. 2767.
15. Bond, P., et al., J Struct Biol, 2007. **157**: p. 593-605.
16. Kandt, C., W. Ash, and D. Tieleman, Methods, 2007. **41**: p. 475.
17. Lemmon, M.A., et al., J. Biol. Chem., 1992. **267**: p. 7683-7689.
18. Treutlein, H., et al., Biochemistry, 1992. **31**: p. 12726-12733.
19. MacKenzie, K.R., J.H. Prestegard, and D.M. Engelman, Science, 1997. **276**(5309): p. 131-3.
20. Smith, S.O., et al., Biophys. J., 2002. **82**: p. 2476.
21. Smith, S.O., et al., Biochemistry, 2001. **40**: p. 6553.
22. Russ, W.P. and D.M. Engelman, J. Mol. Biol., 2000. **296**(3): p. 911-919, doi:10.1006/jmbi.1999.3489.