IET Journals

The Institution of Engineering and Technology

Mechanistic modelling of tyrosine recombination reveals key parameters determining the performance of a CAR T cell switching circuit

Jack E. Bowyer[1,2], Deboki Chakravarti[3,4], Wilson W. Wong[3,4], Declan G. Bates[1,2] ✉

[1]School of Engineering, University of Warwick, Coventry, CV4 7AL, UK
[2]Warwick Integrative Synthetic Biology Centre, Coventry, CV4 7AL, UK
[3]Department of Biomedical Engineering, Boston University, Boston, MA, 02215, USA
[4]Biological Design Center, Boston University, Boston, MA, 02215, USA
✉ E-mail: d.bates@warwick.ac.uk

Abstract: Inducible genetic switches based on tyrosine recombinase-based DNA excision are a promising platform for the regulation and control of chimeric antigen receptor (CAR) T cell activity in cancer immunotherapy. These switches exploit the increased stability of DNA excision in tyrosine recombinases through an inversion–excision circuit design. Here, the authors develop the first mechanistic mathematical model of switching dynamics in tyrosine recombinases and validate it against experimental data through both global optimisation and statistical approximation approaches. Analysis of this model provides guidelines regarding which system parameters are best suited to experimental tuning in order to establish optimal switch performance in vivo. In particular, they find that the switching response can be made significantly faster by increasing the concentration of the inducer drug 4-OHT and/or by using promoters generating higher expression levels of the FlpO recombinase.

1 Introduction

Cell-based therapies that employ engineered T cells expressing chimeric antigen receptors (CARs) to target cancer cells have demonstrated promising responses in recent clinical trials, especially with regard to B cell cancers [1–4]. T cell immunotherapy involves extracting T cells from a patient's blood, engineering the cells to improve their ability to detect and target tumour cells, and reintroducing them back into the patient. Synthetic CARs are integrated into the cells which enable them to recognise specific tumour-associated antigens. There are, however, a number of significant safety and efficacy concerns that have arisen regarding T cell therapies. Success against solid tumours has been limited due to difficulties relating to autoreactive 'on-target, off-tumour' responses as well as cytokine release syndrome, a response to antigen stimulation that accelerates the immune response to potentially fatal levels. Strategies such as kill switches, designed to terminate the T cell response in cases of high toxicity, terminate the entire therapy which often leads to the continuation of tumour development. Reliable modulation of the therapy is therefore of great importance and has prompted research in the development of switching and control circuitry to enable the regulation of engineered T cell responses [5–17].

A promising approach for the creation of such circuitry is the use of site-specific recombinases (SSRs), which are capable of precise DNA manipulation in mammalian cells [18–20], and can be used to engineer synthetic biological circuits capable of performing user-defined functions that are programmed into the cellular DNA [21, 22]. Moreover, SSR-based gene switches elicit permanent alterations to the DNA and thus possess a built-in memory property. As such, SSR-based gene switches are capable of stable changes in gene expression after transient exposure to drug inducer. This memory property could simplify the implementation of such circuitry in gene therapy or cellular immunotherapy, where continuous drug dosing to control transgene expression in patients presents a logistic challenge.

SSRs belong to two main families, the serine and tyrosine recombinases. The former mediate unidirectional (irreversible) recombination events that produce stable genetic states [23, 24], however, the reversal of recombination events via serine integrases is typically dependent on a recombination directionality factor (RDF) [25–28], which can compromise their efficiency [29, 30]. Consequently, serine recombinase-based circuit designs attempt to exploit unidirectional recombination events that are, therefore, unable to facilitate transitioning back to prior system states [31]. Recent mathematical modelling studies regarding serine integrases and their RDFs have revealed that computational models are capable of quantitative replication of the observed in vitro dynamics [30, 32] and can aid the examination of the practical viability of novel synthetic circuits that combine multiple biological elements [30, 33]. Here we seek to develop the first equivalent model for tyrosine recombinase-based circuitry and explore its use in determining key experimental parameters that govern the performance of a CAR T cell switching circuit.

Tyrosine recombinases elicit a primary inversion event in an identical manner to that of serine integrases. The secondary inversion event, however, can also be mediated by the same tyrosine recombinase without any additional cofactors or RDFs, which causes bidirectionality (reversibility) (Fig. 1*a*) [34, 35]. Tyrosine recombinase binding interactions exhibit positive cooperativity [19, 36, 37] and comprise a series of sequential single strand exchanges known as a Holliday junction [38, 39]. Deletion and insertion events mediated by tyrosine recombinases also mirror that of their serine cousins, with the exception of bidirectionality (Fig. 1*b*). Hence, adopting an elementary inversion switch design is not viable, since inversion events are susceptible to natural reversibility, which compromises the efficiency of the desired transcriptional response (Fig. 1*c*). Reversibility is theoretically more prominent in inversion events compared to deletion and insertion events. This is because the deletion of a genetic sequence forms a loop of DNA that dissociates from its original strand, and therefore presents additional spatial disparity to overcome in a subsequent insertion event, compared to successive inversions of the same genetic sequence. Consequently, DNA excision has emerged as a viable method for developing tyrosine recombinase-based circuitry [40, 41]. The inducible control over gene
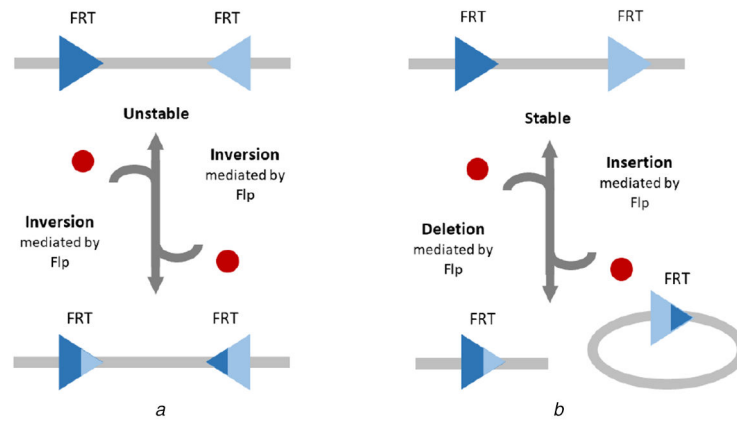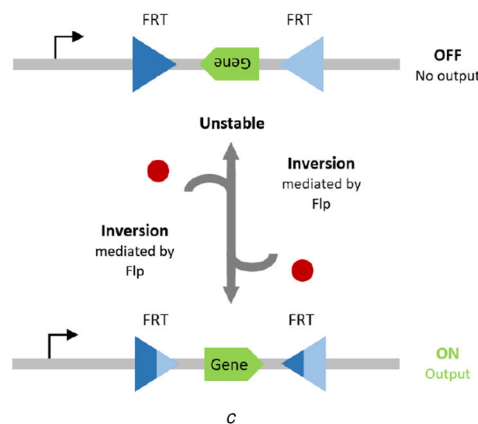
Fig. 1 *Schematic diagrams of DNA recombination events mediated by tyrosine recombinases*
(a) Antiparallel FRT sites result in reversible DNA inversion results, *(b)* Parallel FRT sites result in reversible DNA excision with increased stability, *(c)* Elementary inversion-based tyrosine recombinase circuit design is not viable for reliable control over gene expression

expression afforded by SSRs has the potential to significantly improve the efficacy of engineered T cells. Recombinase-based genetic switches represent the most versatile switches in T cells, which have the potential to enable optimal T cell activity that can be tuned via drug dosage and duration, whilst retaining the memory of transient stimuli.

Mathematical models have been used to elucidate a wide variety of synthetic biological circuits. These include gene regulatory networks [42, 43], tuneable oscillators [44], and genetic counters [21]. Problems that arise as a side effect of assembling synthetic biological circuits can also be addressed using mathematical models such as the burden placed on cell growth due to the introduction of synthetic constructs [45]. At the very least, *in silico* simulation of a given circuit can test hypotheses relating to its dynamic performance in a fraction of the time taken to carry out the equivalent experimental study. The development of new models, therefore, has the potential to create a host of valuable tools for synthetic biologists looking to design and implement novel circuitry in vivo [46]. In this study, we present a mechanistic mathematical model of an inversion–excision switch based on our knowledge of the underlying molecular interactions associated with tyrosine recombinases. The model is formulated through the application of mass action kinetics, thus providing a deterministic output via numerical simulation. The model is validated through two separate data-fitting procedures. The first is a global optimisation technique that minimises error through a 'survival of the fittest' procedure. The second is a statistical technique that employs Bayes' theorem to infer the parameterisation most likely to have produced the data. Our validated model is used to demonstrate an inverse proportionality between 4-OHT concentration and switch response time and identify which model parameters are most sensitive to perturbations and are hence best suited to experimental tuning in vivo.

2 Results and discussion

2.1 Mechanistic model of a tyrosine recombinase inversion–excision switch

The tyrosine recombinase genetic switch developed in [41] exploits the increased stability of DNA excision through a gated inversion–excision circuit design (Fig. 2). The input to this circuit is the flippase O (FlpO) recombinase, an optimised variant of flippase (Flp) with increased efficiency within mammalian cells [47]. A metabolite of the Food and Drug Administration-approved drug tamoxifen, known as 4-OHT, is used to activate the FlpO before it can enter the nucleus and mediate the necessary DNA recombination events. FlpO is expressed constitutively in the cytosol until the advent of 4-OHT causes activation and transport into the nucleus. Leaky nuclear localisation of FlpO can also occur in the absence of 4-OHT, resulting in basal switch activity.

The full network of molecular interactions underlying the switch is shown in Fig. 2, and consists of each DNA:protein and Holliday junction complex arising from FlpO binding interactions. The initial system state (State 1) consists of a blank sequence flanked by antiparallel FRT sites and an inverted CAR gene flanked by parallel FRT sites. This setup does not provide any CAR expression since the non-expressing blank sequence is positioned immediately downstream of the promoter and the CAR gene is not suitably oriented for transcription. FlpO expression elicits inversion of either FRT site pair to produce two transient genetic states. The first sees inversion of both genes, resulting in the CAR expression due to the now suitably oriented CAR gene
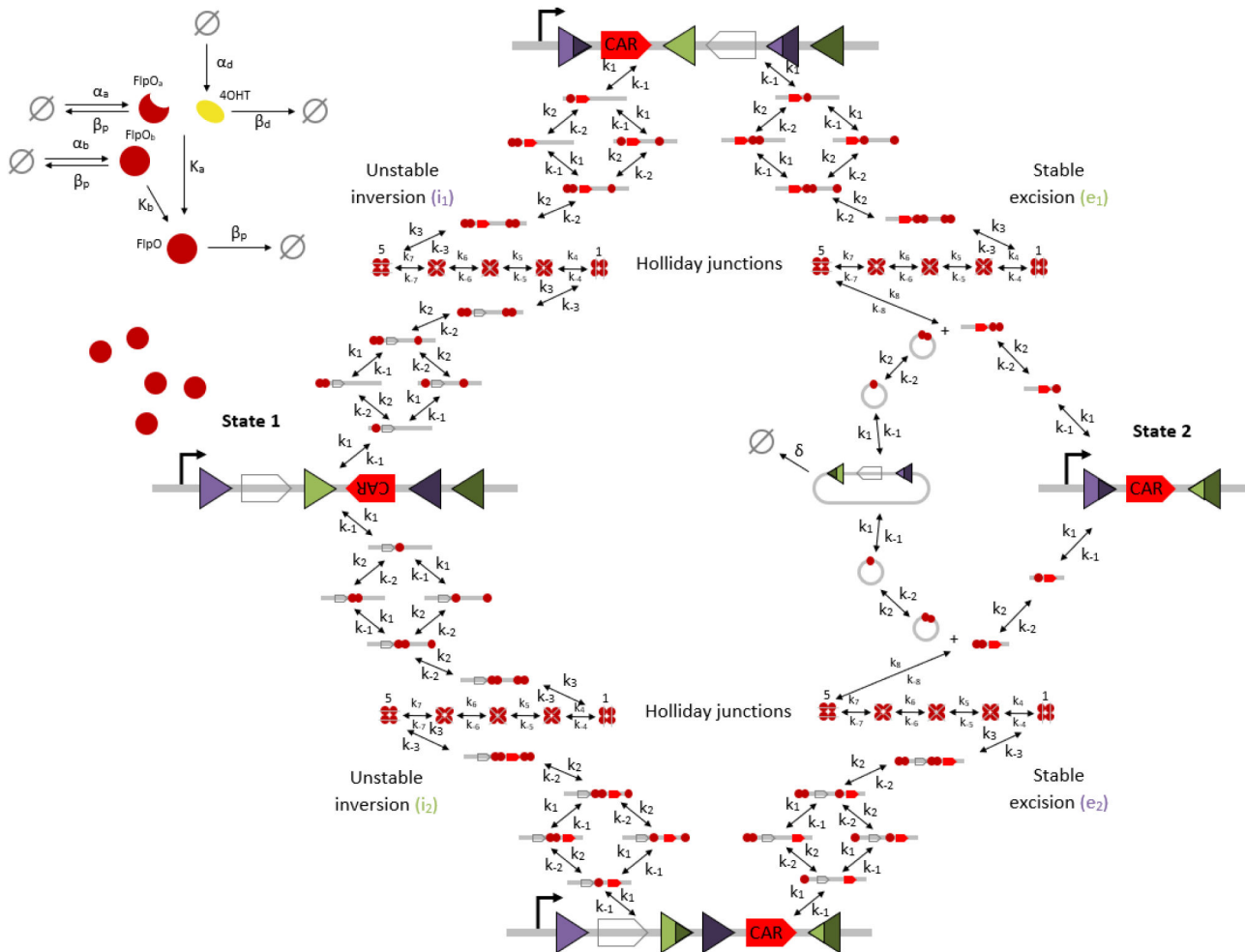
Fig. 2 *Schematic diagram of a tyrosine recombinase-based inversion–excision switch. FlpO recombinase in its activated form targets orthogonal pairs of FRT attachment sites in transitioning the switch from State 1 to State 2 via two transient intermediate genetic states. FRT sites are depicted as purple and green triangles with different shades used to illustrate the result of recombination events. Clear and red pointed boxes depict the blank sequence and CAR gene, respectively. Black arrows depict DNA:protein binding reactions comprising unstable inversion and stable excision events. The rate of unactivated FlpO expression is denoted by α_a; the rate of expression of activated FlpO that leaks into the nucleus is denoted by α_b; the concentration of 4-OHT in the system is modelled as an expression rate denoted by α_d and is an initial condition of the model. The rate of degradation of the recombinase protein is denoted by β_p; the rate of degradation of drug inducer is denoted by β_d. The rate of FlpO activation and nuclear localisation is denoted by K_a; the rate of leaky FlpO nuclear localisation is denoted by K_b. The rate of a FlpO monomer binding reversibly to free FRT sites is denoted by k_1, k_{-1}; due to the cooperativity of FlpO monomer binding, we denote the rate of a FlpO monomer binding reversibly to an occupied FRT site by k_2, k_{-2}. The parameters k_1, k_{-1}, k_2, k_{-2} are equal for each pair of recombination events involving the same pair of attachment sites, i.e. the i_1 and e_2 parameters are equal and the i_2 and e_1 parameters are equal. The rate of Holliday junction formation is denoted by k_3, k_{-3} for the inversions reactions, and by k_3, k_{-3}, k_8, k_{-8} for the excision reactions. Each of the five Holliday junction strand exchanges is denoted by k_4, k_{-4}, k_5, k_{-5}, k_6, k_{-6}, k_7, k_{-7}, respectively. The rate of dilution of the excised DNA (S_X) due to cell division is denoted by δ. The empty set symbol is used to depict expression and degradation reactions*

immediately downstream of the promoter. The second sees the inversion of the CAR gene downstream of the blank sequence which also results in CAR expression. Both transient states consist of the blank sequence flanked by parallel FRT sites, and subsequent FlpO binding, therefore, elicits excision events that delete this sequence. Excision results in a shortened DNA sequence that consists of the CAR gene downstream of the promoter, which therefore provides CAR expression (State 2). In summary, State 1 provides no gene expression, the transient states both provide CAR expression and State 2 also provides CAR expression.

There are four molecular reaction networks comprising the inversion–excision switch system; two unstable inversion networks that detail the transition from State 1 to the two transient genetic states (i_1, i_2), and two stable excision networks that detail the transition from the transient genetic states to State 2 (e_1, e_2). The inversion reaction network is symmetrical due to the fact that successive inversion events are identical. Free DNA is bound by monomeric FlpO until two monomers associate with each of the corresponding attachment sites whereby inversion is initiated. We account for the cooperativity of monomer binding by modelling

two distinct DNA:protein complexes with one FlpO monomer bound at each attachment site $(D_{S_p}^{i_n/e_m}F_{1,1})$ and two monomers bound at one site $(D_{S_p}^{i_n/e_m}F_{2,0})$. Details of the mathematical representation of cooperativity in our model can be found in Section 3. Our model consists of five intermediate Holliday junction complexes describing the inversion/excision of the DNA sequence. FlpO dissociates from the resultant composite attachment sites to give free, genetically differentiated DNA. The excision reaction is not symmetrical, due to the spatial disparity of its products. Again, two FlpO monomers bound at each corresponding FRT site is sufficient to facilitate the excision event, and Holliday junction formation is identical to that of the inversion reaction. FlpO dissociation produces free disparate genetic products: the excised sequence of DNA whose exposed ends are ligated to form a loop, and the shortened DNA strand that remains. Transition to prior genetic states is made possible by the reversibility of the molecular reactions, however, spatial localisation is required for maximum efficiency in the case of transitioning back to State 1 from State 2 (not relevant for the applications presented here).

Although this circuit resembles biological logic gate circuitry in the literature, it has just one input and its only desired function is to implement an efficient ON switch from State 1 to State 2. The ON switch is designed to allow the activation of the CAR T cells to be delayed until an appropriate time point, based on the requirements of individual patients. The OFF switch can be considered as a separate circuit in which the initial positions of the blank sequence and CAR gene are swapped, and thereby only State 1 provides CAR expression. The OFF switch is designed to provide controlled cessation of an initial ON state as an alternative to more traditional kill switches that terminate the therapy completely. Both the ON and OFF switches are induced in the same manner, through 4-OHT activation of the FlpO recombinase which then performs the necessary DNA recombination.

Based on a detailed analysis of the underlying reactions, we developed a set of biochemical equations representing the molecular interaction network shown in Fig. 2 (see supplementary material, Table S1). Our mechanistic model is derived through the application of mass action kinetics to these biochemical equations. The result is a system of ordinary differential equations (ODEs) that describes the rate of change in concentration of each molecular entity with respect to time (see supplementary material, Table S2). Since we are interested in the dynamical response of the system over time and have time course data with which to validate the model, this modelling approach provides suitable approximations to the time-dependent fluctuations in the concentration of model variables in the ON state (the percentage of cells ON). The deterministic output provides an approximation to the average of an ensemble of stochastic simulations without the increased computational workload necessary for such probabilistic approaches.

To simulate the time course concentration dynamics of the genetic states of interest, we compute the total concentration of the system in these states by summing all ODEs describing molecular entities in the same state. Hence, we can derive the following two equations that capture the time course evolution of State 1 and State 2:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = k_3\left(D_{tS_1}^{i_1}F_4 + D_{tS_2}^{i_2}F_4\right) - k_{-3}\left(H_5^{i_1} + H_5^{i_2}\right), \quad (1)$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = k_{-3}\left(H_5^{i_1} + H_5^{i_2}\right) - k_3\left(D_{tS_1}^{i_1}F_4 + D_{tS_2}^{i_2}F_4\right)$$
$$+ \frac{k_8}{2}\left(D_{S_2}^{e_1}F_2 \cdot D_{S_X}^{e_1}F_2 + D_{S_2}^{e_2}F_2 \cdot D_{S_X}^{e_2}F_2\right) \quad (2)$$
$$- \frac{k_{-8}}{2}\left(H_5^{e_1} + H_5^{e_2}\right),$$

where S_1, tS_1, S_2, tS_2 and S_X denote State 1, transient State 1, State 2, transient State 2 and the deleted DNA state respectively; D, H and F denote DNA, Holliday junction complexes and FlpO, respectively; i_1, i_2 and e_1, e_2 denote each of the two inversion and excision reactions, respectively. Numerical subscripts for H and F denote the relevant Holliday junction complex (out of five) and the number of monomers, respectively. The numerical solutions to (1) and (2) constitute the outputs of our model simulations. The factor of half in the latter two terms in (2) is important for preserving the conservation of concentration (mass) in the model. A description of how this factor arises mathematically, using a simplified example, is provided in Supplementary Material, Section S1.

2.2 Model accurately predicts circuit switching efficiency and dynamics

Our experimental data on the performance of the inversion–excision switch consists of four time-course datasets that record the percentage CAR expression of both the ON switch and OFF switch over a ten-day period subject to 0 and 1 μM 4-OHT dosages (see Section 3).

We employed two separate methods to fit the model to our time course data, due to the lack of experimental data regarding the values of our model parameters. The first method employs global

optimisation using a genetic algorithm (GA). The GA mimics natural selection by evolving an initial population of randomly generated solutions over a large number of generations until it converges to a near global minimum within the allocated parameter space (see Section 3). Such algorithms have been widely used in mathematical model inference research relating to, e.g. synthetic oscillators and gene regulatory networks [48–50].

The second data fitting method is a statistical approach known as approximate Bayesian computation (ABC). This is implemented using ABC-SysBio, a Python software package designed specifically for statistical parameter inference in biological systems research [51–53]. The programme employs sequential Monte Carlo (SMC) simulations to construct an accurate approximation to the posterior probability distribution defined by Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad P(B) > 0.$$

Monte Carlo approaches involve computational simulations to generate random candidate solutions, testing their fitness against the desired output and repeating until a viable solution can be identified (see Section 3). ABC establishes distributions on each parameter value in order to infer the parameterisation most likely to have produced the data. In contrast, the GA identifies a single model parameterisation that provides minimal error between the associated in silico simulation and experimental data. However, a single parameterisation provides no information on the extent of uncertainty in the optimal values and is thus more susceptible to overfitting the data. Consequently, we seek maximum reliability in our results by initially performing global optimisation, then verifying this result by examining the likelihood of the existence of better solutions using ABC.

Global optimisation data fitting results confirm that the mechanistic model is capable of replicating time course percentage CAR expression for both ON and OFF switches over the ten-day period recorded experimentally (Fig. 3, solid lines). The optimal parameter values identified by the GA are listed in supplementary material, Table S3. The error function used to assess the fitness of each solution calculates the absolute error across all four datasets. The optimal parameterisation provides an error of 28.15, which enables us to inform the selection of the sequence of error thresholds, ϵ_i, for the subsequent ABC parameter inference. In this case, we set the penultimate error threshold, ϵ_9, at 29 and the final threshold, ϵ_{10}, at 28.15 under the assumption that solutions providing an error below 29 are sufficiently similar to our global optimal solution, and that convergence to the final threshold will present solutions superior to that of the GA. ABC parameter inference results reveal that a population of 100 parameterisations met the penultimate error threshold, but that convergence to the final threshold did not occur within a feasible time frame (Table 1). This provides strong evidence that the parameterisation identified by the GA is indeed near the global optimal solution since we are unable to statistically verify the existence of other solutions capable of providing a smaller error. Furthermore, our ABC approach provides distributions of each parameter value for accepted solutions at each error threshold. These distributions are non-uniform and hence it is unlikely that multiple distinct parameterisations are capable of providing the same optimal model output. This also suggests that our model is identifiable; at least within a finite region of the parameter space.

In order to clarify the key mechanistic features of the system, we conducted a sensitivity analysis to identify whether, and how, the model could be simplified without compromising performance (supplementary material, Fig. S1). Parameters are varied by up to two orders of magnitude above and below their optimal values. Sensitivity is examined with respect to response time, which is measured as the time taken for the ON switch to reach 90% of its maximum percentage CAR expression, and the time taken for the OFF switch to reach 90% of its total decrease in percentage CAR expression.

All parameters describing FlpO expression/degradation mechanisms within the system (α_a, α_b, β_p, β_d, K_a, K_b) showed high sensitivity, and are therefore retained in the simplified model. The
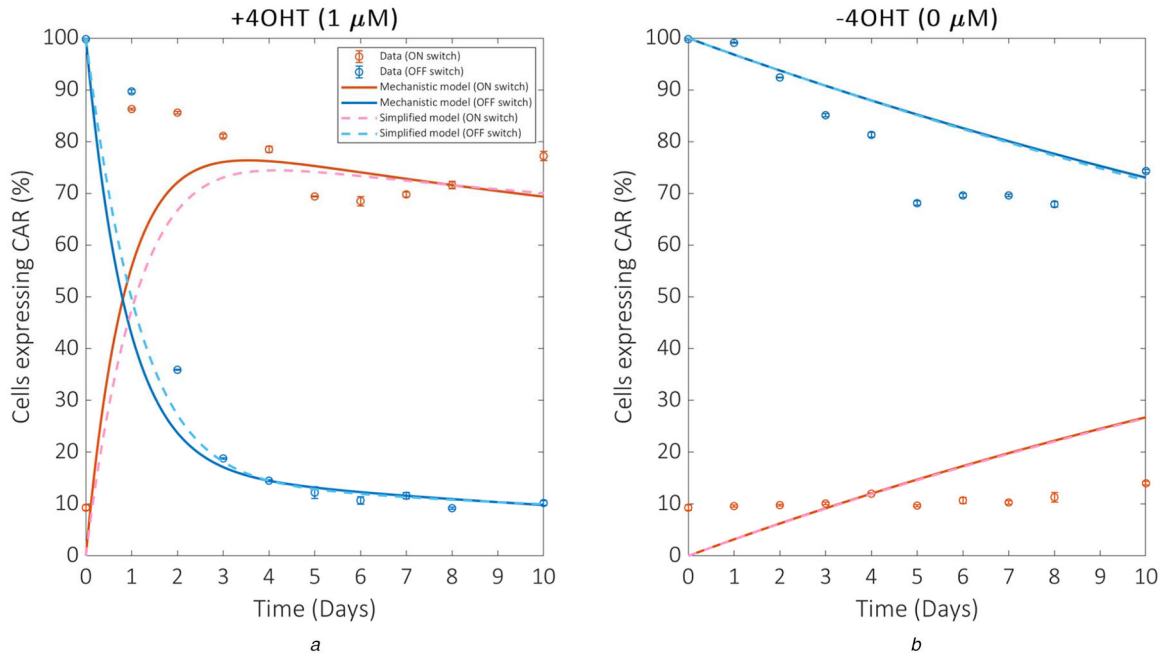
Fig. 3 *Data fitting results. The model is optimised against four time course datasets simultaneously: the ON and OFF switch responses to*
(a) +4-OHT (1 μM 4-OHT) and (b) −4-OHT (0 μM 4-OHT). Red and blue lines depict ON and OFF switch model simulations corresponding to red and blue circles depicting ON and OFF switch experimental data, respectively. Solid lines depict mechanistic model simulations. Dashed lines depict simplified model simulations

Table 1 ABC-SysBio parameter inference results

Error thresholds	Particles assessed	Particles accepted, %	Time, s
$\varepsilon_1 = 500$	100	100.00	227
$\varepsilon_2 = 250$	100	100.00	236
$\varepsilon_3 = 100$	647	15.46	1674
$\varepsilon_4 = 50$	2019	4.95	9031
$\varepsilon_5 = 45$	2736	3.65	12,040
$\varepsilon_6 = 40$	3873	2.58	16,821
$\varepsilon_7 = 35$	8755	1.14	38,438
$\varepsilon_8 = 30$	57,151	0.17	248,534
$\varepsilon_9 = 29$	104,289	0.096	456,383
$\varepsilon_{10} = 28.15$	—	—	—

parameters describing Holliday junction formation (k_4, k_{-4}, k_5, k_{-5}, k_6, k_{-6}, k_7, k_{-7}) showed relatively low sensitivity and are removed. The parameters describing cooperative FlpO monomer binding showed a high sensitivity for one pair of attachment sites (i_2, e_1: k_1, k_{-1}, k_2, k_{-2}), but comparatively low sensitivity for the other pair (i_1, e_2: k_1, k_{-1}, k_2, k_{-2}). Since each of the two pathways from State 1 to State 2 features recombination events involving each pair of attachment sites, we remove cooperative monomer binding and describe each recombination event as a single reversible reaction. The reduced model retains the four state variables relating to FlpO expression, activation, and degradation in the mechanistic model, but consists of just five DNA state variables compared to 63, while the number of parameters is reduced from 27 to 12 (Fig. 4). Applying mass action kinetics to the biochemical reactions depicted in Fig. 4 gives the following system of nine ODEs:

$$\frac{d[F]}{dt} = K_a \bar{V}[4OHT][F_a] + K_b \bar{V}[F_b] - \beta_p[F] - k_1[F][D_{S_1}]$$
$$- k_r[F][D_{tS_1}] - k_2[F][D_{S_1}] - k_r[F][D_{tS_2}]$$
$$- k_3[F][D_{tS_1}] - k_4[F][D_{tS_2}], \tag{3}$$

$$\frac{d[4OHT]}{dt} = \alpha_d - \beta_d[4OHT] - K_a[4OHT][F_a], \tag{4}$$

$$\frac{d[F_a]}{dt} = \alpha_a - \beta_p[F_a] - K_a[4OHT][F_a], \tag{5}$$

$$\frac{d[F_b]}{dt} = \alpha_b - \beta_p[F_b] - K_b[F_b], \tag{6}$$

$$\frac{d[D_{S_1}]}{dt} = k_r[F][D_{tS_1}] - k_1[F][D_{S_1}] + k_r[F][D_{tS_2}]$$
$$- k_2[F][D_{S_1}], \tag{7}$$

$$\frac{d[D_{tS_1}]}{dt} = k_1[F][D_{S_1}] - k_r[F][D_{tS_1}] - k_3[F][D_{tS_1}]$$
$$+ k_r[D_{S_X}][D_{S_2}], \tag{8}$$

$$\frac{d[D_{tS_2}]}{dt} = k_2[F][D_{S_1}] - k_r[F][D_{tS_2}] - k_4[F][D_{tS_2}]$$
$$+ k_r[D_{S_X}][D_{S_2}], \tag{9}$$

$$\frac{d[D_{S_X}]}{dt} = k_3[F][D_{tS_1}] - k_r[D_{S_X}][D_{S_2}] + k_4[F][D_{tS_2}]$$
$$- k_r[D_{S_X}][D_{S_2}] - \delta[D_{S_X}], \tag{10}$$

Each population of accepted particles is formed by assessing the error provided by each particle. If the error is less than the threshold for that population the particle is accepted if not, the particle is rejected. Our inference required that 100 particles be accepted before progression to the next population and error threshold. The final error threshold, ε_{10}, was chosen based on our GA optimisation result and ABC-SysBio was unable to identify a population of particles that provide less error than this result

$$\frac{d[D_{S_2}]}{dt} = k_3[F][D_{tS_1}] - k_r[D_{S_X}][D_{S_2}] + k_4[F][D_{tS_2}]$$
$$- k_r[D_{S_X}][D_{S_2}]. \tag{11}$$

To assess the performance of the simplified model in reproducing the experimental data we again implemented global optimisation. GA data fitting results reveal that the simplified model is capable of matching the minimal error provided by the optimal mechanistic
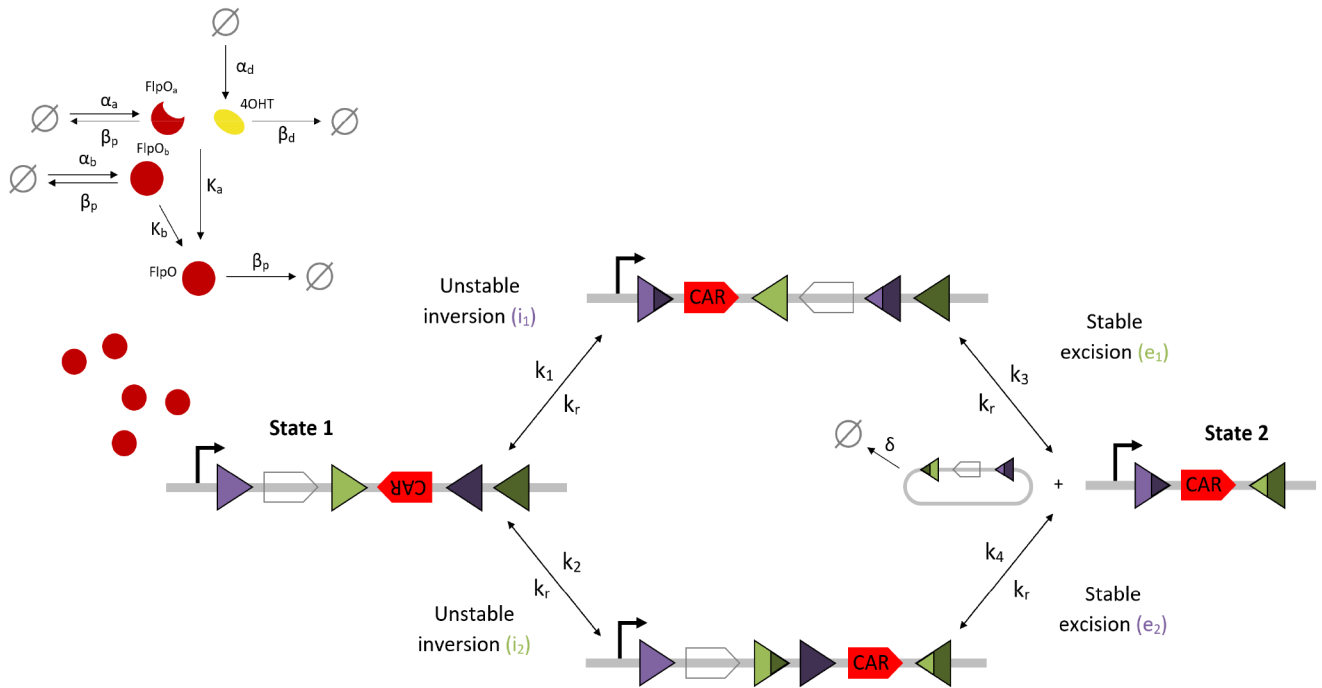
Fig. 4 *Schematic diagram of the reduced tyrosine recombinase-based inversion–excision switch. FlpO recombinase, in both activated (FlpO) and inactivated (FlpOa, FlpOb) forms, targets gated FRT sites (purple and green triangles) in transitioning the switch from State 1 to State 2 via two transient intermediate genetic states. Clear and red pentagons depict the blank sequence and CAR gene respectively. Black arrows depict DNA:protein binding reactions comprising unstable inversion and stable excision events. Reaction rate constants are denoted by the corresponding numbered k. The parameter k_r denotes the rate of the reverse reaction for each recombination which enables further reduction of the number of model parameters*
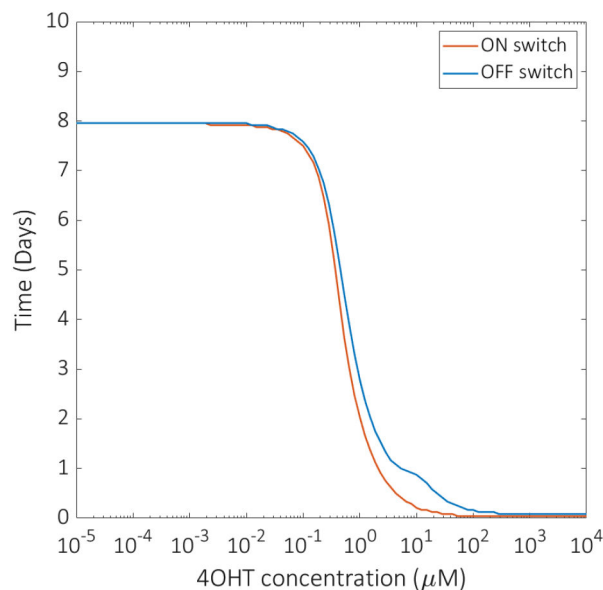

Fig. 5 *Model predictions of drug dosage dependency. Optimised reduced model simulations of response time as a function of 4-OHT concentration for both the ON switch (red line) and OFF switch (blue line). Response time is measured as the time taken for the ON switch to reach 90% of its maximum percentage CAR expression, and the time taken for the OFF switch to reach 90% of its total decrease in percentage CAR expression*

model parameterisation (Fig. 3, dashed lines). The optimal parameterisation for the reduced model provides an error of 28.39, a negligible increase on the 28.15 provided by the mechanistic model. The optimal parameter values for the reduced model are listed in supplementary material, Table S4. Further attempts to simplify the model by removing other features resulted in large increases in the fitting error, providing strong evidence that this model represents the simplest possible mechanistic representation of the system. All results described hereafter were derived from this simplified model of the system.

2.3 Switching response times are inversely proportional to the concentration of inducer drug

Response time is the key performance characteristic of the switch since ultimately it will need to function reliably in mammalian cells within time frames that are appropriate to the condition of individual patients. By identifying parameters that are both influential over system output and experimentally tuneable, we can establish guidelines to optimise the functioning of the switch in different application scenarios. Model simulations reveal that there is an inverse proportionality between drug concentration and switching response time (Fig. 5). The response time of both switches is significantly decreased when induced by 4-OHT concentrations in the range 10^{-1}–10^1 μM. Therefore the 1 μM 4-
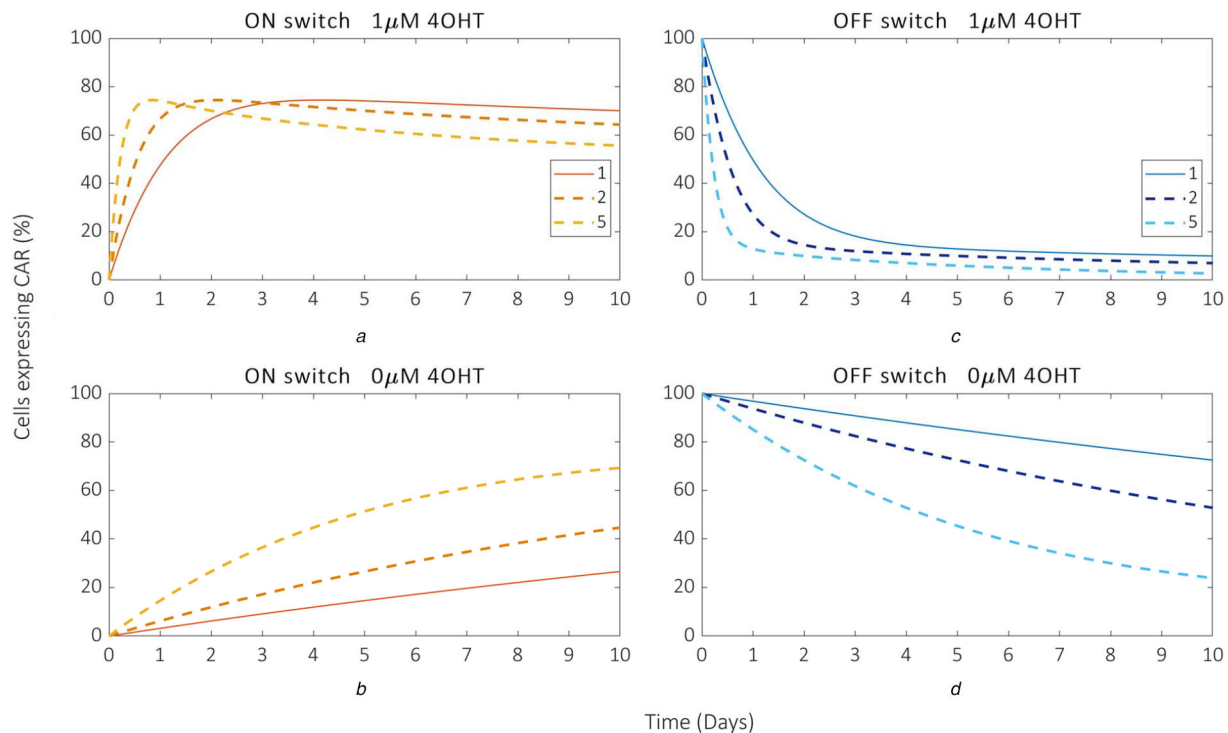
Fig. 6 *Model predictions of the effect of tuning FlpO expression on switch response times. Model simulations of ON switch and OFF switch time course responses for the optimal values of α_a and α_b subject to fold changes of one, two and five*

(*a*) ON switch + 4-OHT (1 μM), (*b*) ON switch − 4-OHT (0 μM), (*c*) OFF switch + 4-OHT (1 μM), (*d*) OFF switch − 4-OHT (0 μM)

OHT concentration used experimentally may require an increase of up to 10 μM to deliver improved response times. The concentration of 4-OHT used to induce the switch is important as it has implications regarding the cost of the treatment. Identifying key operational factors through model simulations ultimately informs the direction of confirmatory experimental work and presents new avenues of investigation in a time efficient manner. Given that the treatment is likely to benefit from minimising the concentration of the inducer drug, the model can be utilised to identify alternative parameters that have the potential to improve switch response time.

2.4 Increasing Flp0 expression can provide faster switching

Sensitivity analysis of the reduced model parameters (supplementary material, Fig. S2) identifies the most sensitive to be α_a, β_p, β_d, and K_a. Of these, the rates of protein and drug degradation, β_p and β_d, and the rate of FlpO nuclear localisation through 4-OHT activation, K_a, are challenging candidates for experimental tuning. However, the rate of FlpO expression in the system, α_a and α_b is potentially tuneable through appropriate choice of the cognate promoter strength. Model simulations reveal faster response times for both the ON and OFF switches as the values of α_a and α_b are increased two- and five-fold (Fig. 6), with a significant reduction in switch response time for even a two-fold increase in FlpO expression. Recent work on engineering Jurkat T cells with inducible promoters of varying strengths illustrates the potential to provide the required increase in FlpO expression. However, the model predicts that increased FlpO expression also results in a decreased percentage of CAR expression over time with respect to the ON switch (Figs. 6*a* and *b*). For example, a five-fold increase in FlpO expression is predicted to provide >75% cells expressing CAR within a relatively small time frame (~1 day), but this reduces to ~55% after 2 days. Therefore, unless the action of the CAR T cells is as effective as to provide significant therapeutic benefit over small time periods, these dynamics are unlikely to be advantageous. By contrast, the performance of the OFF switch improves uniformly as FlpO expression increases.

The model also predicts that increased FlpO expression causes increased basal switching responses from both switches (Figs. 6*c* and *d*) that could potentially compromise the efficacy of the therapy altogether, regardless of which switch was considered to be most viable. Hence, attempts to improve the efficiency of the switch through increased FlpO expression would not be suitable for applications in which tight regulation of the initial state in the absence of induction is required.

2.5 Optimal switch performance presents a trade-off between 4-OHT concentration and FlpO expression

Investigating the effects of simultaneously modifying the expression of FlpO and the concentration of 4-OHT reveals that improved switch response times are possible for relatively large doses of 4-OHT if FlpO expression is unaltered, and for relatively low doses of 4-OHT if FlpO expression is increased ~10-fold (Fig. 7). The model predicts that increasing FlpO expression ~2-fold and using a 4-OHT dosage of ~3 μM can provide a notable improvement in response times for both the ON and OFF switches. Although this demonstrates an advantageous reduction in the magnitude of the required fold change in FlpO expression for a relatively small increase in drug dosage, the resulting basal activity increases to potentially problematic levels, suggesting that tuning inducer concentration/dosage may be a more viable method for improving switch efficiency.

3 Materials and methods

3.1 Experimental procedure

Primary CD4 + T cells were harvested from blood obtained from the Boston Children's Hospital using the STEMCELL CD4 + enriched cocktail and RosetteSep system. T cells were maintained in X-VIVO 15 media (Lonza) supplemented with 5% AB Serum (Valley Biomedical), 10 mM *N*-acetyl L-cystein (Sigma), 55 μM 2-mercaptoethanol (Gibco), and 50–100 units/ml recombinant interleukin-2 (IL-2) (Tecin, NCI BRB Preclinical Repository).

Lentivirus containing either the inducible FlpO or stable inversion switch was produced via polyethylenimine transfection of the human embryonic kidney 293 FT cells and collected 3 and 4 days post-transfection. The concentrated virus was produced via ultracentrifugation with 20% sucrose (Sigma) for 2 h at 4°C and 22,000 *g*. T cells were activated with CD3/CD28 Dynabeads (Gibco), and then spinfected with a virus one day later. Using half
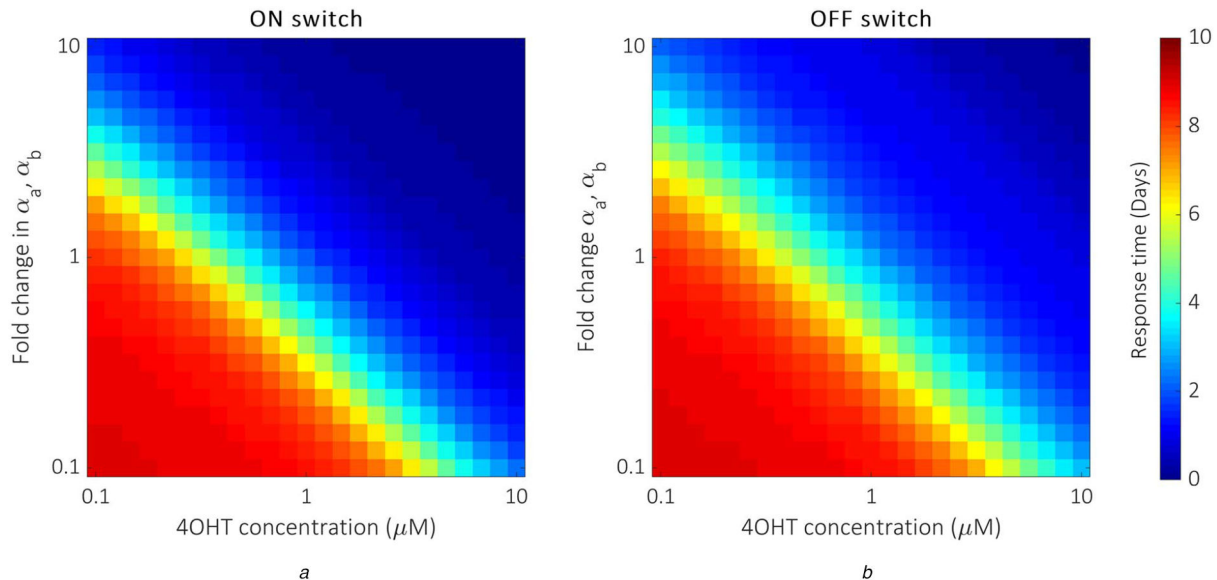
Fig. 7 *Model predictions of parameter tuning trade-offs*

(*a*) Model simulations of ON switch response time for the optimal values of α_a and α_b subject to fold changes of one order of magnitude above and below the optimal values, and for 4-OHT concentrations between 10^{-1} and 10^1 μM, (*b*) Model simulations of OFF switch response time for the optimal values of α_a and α_b subject to fold changes of one order of magnitude above and below the optimal values, and for 4-OHT concentrations between 10^{-1} and 10^1 μM. Response time is measured as the time taken for the ON switch to reach 90% of its maximum percentage CAR expression, and the time taken for the OFF switch to reach 90% of its total decrease in percentage CAR expression. The colour gradient ranges from dark blue (0 days) to dark red (10 days)

of the concentrated virus, both inducible recombinase and stable inversion switch viruses were spun onto the well a six-well retronectin (Clontech)-coated plate for 90 min at 1200 *g* and activated CD4 + T cells were then spun on to the virus plates for 60 min at 1200 *g*. T cells were sorted using the SH800 Cell Sorter (Sony, BFP-FL1 channel) for FlpOER^{T2} expression using a BFP marker expressed in the FlpOER^{T2} cassette via a T2a ribosomal skip sequence.

Cells were induced with 1 μM 4-hydroxytamoxifen (4-OHT, Sigma) in methanol solution except for dose response experiments. Cells were induced at a starting concentration of 200,000 cells/ml and maintained between 200,000–1,200,000 cells/ml with media containing 4-OHT. Uninduced cells were also plated and maintained at the same cell concentrations in inducer-free media. CAR expression was observed via phycoerythrin (PE)-conjugated antibody staining (R&D Systems IC3696P) for a human myc epitope tag expressed in the extracellular portion of the CAR, which was then quantified via flow cytometry (Attune NxT Flow Cytometer).

The percentage switching efficiency is thought to reflect the efficiency of integrating both the FlpO-producing component and the switch itself into the T cells. Given that successful integration is not guaranteed for every cell, there are four possible outcomes: R−/FS− (no components), R+/FS− (recombinase only), R−/FS+ (switch only) and R+/FS+ (recombinase and switch). Only cells that have acquired the requisite components (R+/FS+) can provide the desired function, and hence we convert our data to represent percentage CAR expression with respect to the global maximum data value to focus our analysis purely on the R+/FS+ cells in the population. We assume that integration efficiency does not vary significantly across experimental trials, allowing us to use the global maximum data value as a metric for functionally viable T cell populations.

All cells are thought to possess an initial volumetric ratio (the ratio of the volume of the cytoplasm to the volume of the nucleus, $V_C : V_N$) of ~4:1 that changes with cell growth, becoming 2:1 or 1:1. However, lymphocytes such as T cells have not been found to demonstrate these growth related adaptations, maintaining a ratio between 4:1 and 3:1. Hence, the parameter $\bar{V}$, equal to the quotient V_C/V_N, is fixed at 0.3 in our model.

We imposed the conditions $k_1 < k_2$ and $k_{-1} > k_{-2}$ to account for cooperative FlpO monomeric binding to DNA, whereby the binding affinity of a FlpO monomer to an attachment site already bound by another monomer is greater than that of a monomer to a free site. This involves setting $k_1 = k_2$ and $k_{-1} = k_{-2}$ and introducing the multipliers $0 < m_1 < 1$ and $0 < m_2 < 1$ such that m_1 multiplies k_1 and m_2 multiplies k_{-2}, hence ensuring that $k_1/k_{-1} < k_2/k_{-2}$ in our parameter inference.

3.2 Model parameter inference and simulation

3.2.1 Global optimisation: We employed the built-in GA function in MATLAB for parameter inference purposes. The GA proceeds by initially generating a population of solutions at random within the predefined parameter space. The initial solutions are then scored based on their fitness. This is typically calculated by virtue of an error function that determines how well each solution is able to match the relevant experimental data. The best solutions are selected as parents that will produce the best offspring to populate the next generation of solutions. Parents produce offspring through crossover, whereby a random place in their binary genotype is selected and the information beyond that point is swapped over. The mutation is also incorporated, whereby single point alterations in the offspring's genotype are imposed to increase diversity within the population. This procedure is repeated indefinitely until a predefined termination criterion is reached; typically, the scores are unchanged over a specified number of generations, the algorithm reaches a specified number of generations, or a specified solution score is established. Parallelisation of the MATLAB code enabled us to run the GA with a large search population for many generations; this significantly increases the likelihood of establishing the global optimal solution. Given the lack of available data regarding the relevant reaction rate constants in the literature, we require a parameter space large enough to locate optimal solutions, but not so big that convergence timescales become impractical. A large parameter space increases the time taken to produce the next generation of solutions however, this can also decrease the overall number of generations required for convergence, since more solutions are inspected in each case. Large parameter spaces also increase the risk of incurring problems with stiff model simulations that may cause the GA to fail, hence multiple trials are often required to determine effective performance criteria. Hence, the search interval imposed on all model parameters is [0, 1]. We ran the GA over 1000 generations with a population size of at least 1000 in order to maximise the likelihood of convergence and the identification of the global

optimum solution. We define the error function that calculates the mean absolute error between the four time-course datasets and their corresponding model simulations such that

$$d = \frac{1}{10} \sum_{i=1}^{40} |x_i - y_i|,$$

where x_i and y_i are corresponding elements of the experimental data and model simulation vectors, respectively. Each of the four time-course datasets consists of ten values and our distance function calculates the mean absolute error across all four datasets.

3.2.2 Approximate Bayesian computation (ABC): The ABC–SMC procedure implemented by ABC-SysBio proceeds in the following manner: an initial population of solutions (particles) is randomly generated in accordance with the prior distributions imposed on the model parameters. Each particle provides a simulated dataset, $D^\star$, which is compared to the fixed experimental dataset, D, by an appropriate distance (error) function and its fitness is scored accordingly. This error score determines the acceptance of a particle, dependent on a decreasing sequence of error thresholds, ϵ, set to correspond with each population, i.e.

$$d(D^\star, D) < \epsilon_i,$$

where $\varepsilon_1 > \varepsilon_2 > \ldots > \varepsilon_n$ and d is the distance function. We define our distance function as the mean absolute error between model simulations and experimental data in the same way as the GA approach. Subsequent populations are obtained by perturbing particles from the previous population in accordance with a predetermined perturbation kernel, proceeding until the model is unable to produce particles of sufficient fitness to satisfy the immediate threshold.

An array of model-specific criteria are required to allow the ABC-SysBio package to run efficiently: the sequence of decreasing error thresholds, ϵ, must be provided whereby only the particles capable of providing error less than that of the threshold will be accepted by the algorithm. Each ϵ must be satisfied in succession until the particles are unable to satisfy the next threshold. The satisfaction of an individual threshold is dependent on the number of particles accepted; the number of acceptable particles required before passage to the next threshold must also be predetermined. The larger the number of particles, the higher probability of reliable inference results and the longer the time taken by the algorithm to converge. Each individual parameter subject to inference requires a prior probability distribution in order to establish the parameter space within which to locate acceptable particles. We allocate the same parameter space used in GA parameter inference to facilitate direct comparisons of the results. Hence, the prior distributions imposed on all model parameters are uniform distributions on the interval [0, 1]. The convergence of the algorithm is dependent on all of the aforementioned factors and hence it may require several trials to establish the appropriate performance criteria. To achieve credible results, it is advised that parameter inference and model selection tasks are repeated multiple times due to the random nature of the Monte Carlo simulations that drive the algorithm.

4 Conclusion

Mathematical models of synthetic circuits that are based strictly on underlying biological mechanisms can be used as design tools to reduce development times, lower the risk of circuit failure, and identify key parameters determining circuit performance. We have developed the first mechanistic mathematical model of a tyrosine recombinase-based inversion–excision switch and validated it against dynamic data on CAR expression using both global optimisation and statistical approaches. A sensitivity analysis of this model revealed key mechanistic features that allowed the construction of a significantly simplified model that yielded equivalent data fitting results. Analysis of this model revealed that the response time of the switch can be modulated experimentally

by tuning the concentration of the inducer drug and/or by choosing promoters that increase the level of expression of the FlpO recombinase. However, both of these solutions also have potential drawbacks in terms of increasing basal expression activity, indicating that the successful clinical deployment of this and other novel recombinase-based circuits might be dependent on the development of new mechanisms for minimising or even nullifying basal recombinase expression.

6 References

[1] Davila, M.L., Riviere, I., Wang, X., *et al.*: 'Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia', *Sci. Transl. Med.*, 2014, **6**, p. 224ra225
[2] Brentjens, R.J., Riviere, I., Park, J.H., *et al.*: 'Safety and persistence of adoptively transferred autologous CD19-targeted T cells in patients with relapsed or chemotherapy refractory B-cell leukemias', *Blood*, 2011, **118**, pp. 4817–4828
[3] Maude, S.L., Frey, N., Shaw, P.A., *et al.*: 'Chimeric antigen receptor T cells for sustained remissions in leukemia', *N. Engl. J. Med.*, 2014, **371**, pp. 1507–1517
[4] Lee, D.W., Kochenderfer, J.N., Stetler-Stevenson, M., *et al.*: 'T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial', *Lancet*, 2015, **385**, pp. 517–528
[5] Wu, C.Y., Roybal, K.T., Puchner, E.M., *et al.*: 'Remote control of therapeutic T cells through a small molecule-gated chimeric receptor', *Science*, 2015, **350**, p. aab4077
[6] Fedorov, V.D., Sadelain, M., Kloss, C.C.: 'Novel approaches to enhance the specificity and safety of engineered T cells', *Cancer J.*, 2014, **20**, pp. 160–165
[7] Kloss, C.C., Condomines, M., Cartellieri, M., *et al.*: 'Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells', *Nat. Biotechnol.*, 2013, **31**, pp. 71–75
[8] Sakemura, R., Terakura, S., Watanabe, K., *et al.*: 'A Tet-On inducible system for controlling CD19-chimeric antigen receptor expression upon drug administration', *Cancer Immunol. Res.*, 2016, **4**, pp. 658–668
[9] Di Stasi, A., Tey, S.K., Dotti, G., *et al.*: 'Inducible apoptosis as a safety switch for adoptive cell therapy', *N. Engl. J. Med.*, 2011, **365**, pp. 1673–1683
[10] Straathof, K.C., Pulè, M.A., Yotnda, P., *et al.*: 'An inducible caspase 9 safety switch for T-cell therapy', *Blood*, 2005, **105**, pp. 4247–4254
[11] Budde, L.E., Berger, C., Lin, Y., *et al.*: 'Combining a CD20 chimeric antigen receptor and an inducible caspase 9 suicide switch to improve the efficacy and safety of T cell adoptive immunotherapy for lymphoma', *PLOS One*, 2013, **8**, p. e82742
[12] Wei, P., Wong, W., Park, J., *et al.*: 'Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells', *Nature*, 2012, **488**, pp. 384–388
[13] Ma, J.S., Kim, J.Y., Kazane, S.A., *et al.*: 'Versatile strategy for controlling the specificity and activity of engineered T cells', *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, pp. E450–E458
[14] Rodgers, D.T., Mazagova, M., Hampton, E.N., *et al.*: 'Switch-mediated activation and retargeting of CAR-T cells for B-cell malignancies', *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, pp. E459–E468
[15] Kudo, K., Imai, C., Lorenzini, P., *et al.*: 'T lymphocytes expressing a CD16 signaling receptor exert antibody-dependent cancer cell killing', *Cancer Res.*, 2014, **74**, pp. 93–103
[16] Chen, Y.Y., Jensen, M.C., Smolke, C.D.: 'Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems', *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, pp. 8531–8536
[17] Torikai, H., Reik, A., Liu, P.Q., *et al.*: 'A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR', *Blood*, 2012, **119**, pp. 5697–5705
[18] Groth, A.C., Olivares, E.C., Thyagarajan, B., *et al.*: 'A phage integrase directs efficient site-specific integration in human cells', *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, pp. 5995–6000
[19] Ghosh, K., Van Duyne, G.D.: 'Cre-loxP biochemistry', *Methods*, 2002, **28**, pp. 374–383
[20] Ghosh, K., Lau, C.K., Guo, F., *et al.*: 'Peptide trapping of the Holliday junction intermediate in Cre-loxP site-specific recombination', *J. Biol. Chem.*, 2005, **280**, pp. 8290–8299
[21] Friedland, A.E., Lu, T.K., Wang, X., *et al.*: 'Synthetic gene networks that count', *Science*, 2009, **234**, pp. 1199–1202
[22] Fernandez-Rodriguez, J., Yang, L., Gorochowski, T., *et al.*: 'Memory and combinatorial logic based on DNA inversions: dynamics and evolutionary stability', *ACS Synth Biol.*, 2015, **4**, pp. 1361–1372
[23] Groth, A., Calos, M.: 'Phage integrases: biology and applications', *J. Mol. Biol.*, 2004, **335**, pp. 667–678
[24] Fogg, P., Colloms, S., Rosser, S., *et al.*: 'New applications for phage integrases', *J. Mol. Biol.*, 2014, **426**, pp. 2703–2716

[25] Ghosh, P., Pannunzio, N., Hatfull, G.: 'Synapsis in phage Bxb1 integration: selection mechanism for the correct pair of recombination sites', *J. Mol. Biol.*, 2005, **349**, pp. 331–348

[26] Ghosh, P., Bibb, L., Hatfull, G.: 'Two-step site selection for serine-integrase-mediated excision: DNA-directed integrase conformation and central dinucleotide proofreading', *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, pp. 3238–3243

[27] Khaleel, T., Younger, E., McEwan, A., *et al.*: 'A phage protein that binds φC31 integrase to switch its directionality', *Mol. Microbiol.*, 2011, **80**, pp. 1450–1463

[28] Olorunniji, F., Buck, D., Colloms, S., *et al.*: 'Gated rotation mechanism of site-specific recombination by φC31 integrase', *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, pp. 19661–19666

[29] Bonnet, J., Subsoontorn, P., Endy, D.: 'Rewritable digital data storage in live cells via engineered control of recombination directionality', *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, pp. 8884–8889

[30] Bowyer, J., Zhao, J., Subsoontorn, P., *et al.*: 'Mechanistic modeling of a rewritable recombinase addressable data module', *IEEE Trans. Biomed. Circuits Syst.*, 2016, **10**, pp. 1161–1170

[31] Hsiao, V., Hori, Y., Rothemund, P., *et al.*: 'A population-based temporal logic gate for timing and recording chemical events', *Mol. Syst. Biol.*, 2016, **12**, p. 869

[32] Pokhilko, A., Zhao, J., Stark, W.M., *et al.*: 'A simplified mathematical model of directional DNA site-specific recombination by serine integrases', *J. R. Soc. Interface*, 2017, **14**, p. 20160618, doi: 10.1098/rsif.2016.0618

[33] Pokhilko, A., Ebenhoh, O., Stark, W.M., *et al.*: 'Mathematical model of a serine integrase-controlled toggle switch with a single input', *J. R. Soc. Interface*, 2018, **15**, p. 20180160, doi: 10.1098/rsif.2018.0160

[34] Sorrell, D.A., Kolb, A.F.: 'Targeted modification of mammalian genomes', *Biotechnol. Adv.*, 2005, **23**, pp. 431–469

[35] Grindley, N.D., Whiteson, K.L., Rice, P.A.: 'Mechanisms of site-specific recombination', *Annu. Rev. Biochem.*, 2006, **75**, pp. 567–605

[36] Guo, F., Gopaul, D.N., Van Duyne, G.D.: 'Asymmetric DNA bending in the Cre-loxP site-specific recombination synapse', *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, pp. 7143–7148

[37] Van Duyne, G.D.: 'A structural view of Cre-loxP site-specific recombination', *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, pp. 87–104

[38] Chen, Y., Rice, P.A.: 'New insight into site-specific recombination from Flp recombinase-DNA structures', *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **47**, pp. 122–131

[39] Rashel, M., Uchiyama, J., Ujihara, T., *et al.*: 'A novel site-specific recombination system derived from bacteriophage φMR11', *Biochem. Biophys. Res. Commun.*, 2008, **368**, pp. 192–198

[40] Weinberg, B.H., Hang Pham, N.T., Caraballo, L.D., *et al.*: 'Large scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells', *Nat. Biotechnol.*, 2017, **35**, pp. 453–462

[41] Chakravarti, D., Caraballo, L.D., Weinberg, B., *et al.*: 'Inducible gene switches with memory in human T cells for cellular immunotherapy', *ACS Synth. Biol.*, 2019, **8**, pp. 1744–1754

[42] Hancock, E.J., Stan, G.-B., Arpino, J.A.J., *et al.*: 'Simplified mechanistic models of gene regulation for analysis and design', *J. R. Soc. Interface*, 2015, **12**, doi: 10.1098/rsif.2015.0312

[43] Herath, N., Del Vecchio, D.: 'Deterministic-like model reduction for a class of multiscale stochastic differential equations with application to biomolecular systems', *IEEE Trans. Autom. Control*, 2018, **64**, pp. 351–358

[44] Tigges, M., Marquez-Lago, T.T., Stelling, J., *et al.*: 'A tunable synthetic mammalian oscillator', *Nature*, 2009, **457**, pp. 309–312

[45] Nystrom, A., Papachristodoulou, A., Angel, A.: 'A dynamic model of resource allocation in response to the presence of a synthetic construct', *ACS Synth. Biol.*, 2018, **7**, pp. 1201–1210

[46] Zheng, Y., Sriram, G.: 'Mathematical modeling: bridging the gap between concept and realisation in synthetic biology', *J. Biomed. Biotechnol.*, 2010, **2010**, doi: 10.1155/2010/541609

[47] Kranz, A., Fu, J., Duerschke, K., *et al.*: 'An improved Flp deleter mouse in C57Bl/6 based on FlpO recombinase', *Genesis*, 2010, **48**, pp. 512–520

[48] Chen, B-S., Chen, P-W.: 'GA-based design algorithms for the robust synthetic genetic oscillators with prescribed amplitude, period and phase', *Gene. Regul. Syst. Bio.*, 2010, **4**, pp. 35–52

[49] Sirbu, A., Ruskin, H.J., Crane, M.: 'Comparison of evolutionary algorithms in gene regulatory network model inference', *BMC Bioinformat.*, 2010, **11**, p. 59

[50] Smith, R.W., van Sluijs, B., Fleck, C.: 'Designing synthetic oscillatory networks in silico: a generalised evolutionary algorithm approach', *BMC Syst. Biol.*, 2017, **11**, p. 118

[51] Liepe, J., Kirk, P., Filippi, S., *et al.*: 'A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation', *Nat. Protoc.*, 2014, **9**, pp. 439–456

[52] Stumpf, M.: 'Approximate Bayesian inference for complex ecosystems', *F1000Prime Rep.*, 2014, **6**, p. 60

[53] Smith, R., Grohn, Y.: 'Use of approximate Bayesian computation to assess and fit models of *Mycobacterium leprae* to predict outcomes of the Brazilian control program', *PLOS One*, 2015, **10**, p. e0129535