# Evolutionary computation for the design of a stochastic switch for synthetic genetic circuits

Jennifer S. Hallinan, Goksel Misirli, and Anil Wipat

***Abstract*—Biological systems are inherently stochastic, a fact which is often ignored when simulating genetic circuits. Synthetic biology aims to design genetic circuits *de novo*, and cannot therefore afford to ignore the effects of stochastic behavior. Since computational design tools will be essential for large-scale synthetic biology, it is important to develop an understanding of the role of stochasticity in molecular biology, and incorporate this understanding into computational tools for genetic circuit design. We report upon an investigation into the combination of evolutionary algorithms and stochastic simulation for genetic circuit design, to design regulatory systems based on the *Bacillus subtilis sin* operon.**

## I. Introduction

STOCHASTICITY is ubiquitous in biology. At every scale of organization, from the molecule to the ecosystem, organisms must be robust in the face of noise generated both internally and externally. Although noise has long been recognized as a challenge for biological systems to overcome [1], there is growing recognition that noise may be useful, even fundamental, to the successful functioning of living systems [2 - 4]. Stochasticity is therefore a significant issue for synthetic biology, a discipline which aims to apply engineering principles to the design and construction of biological systems with novel, valuable functionality [5]

At present, most synthetic biology projects deal with relatively small numbers of genes, operating in restricted sets of pathways. However, the ultimate aim of the synthetic biology endeavor is to design large-scale genetic circuits, incorporating components which originally evolved in a range of species. Over the last few years considerable progress has been made towards engineering fully synthetic organisms [6, 7].

Small systems can be designed by hand, and their performance tweaked on the basis of observation, making current synthetic biology an essentially empirical discipline [8]. However, the large-scale design of genetic circuits will require computer-aided design (CAD) and simulation tools, which should produce circuit designs with behavior as predictable as those of electrical circuits designed by electrical engineering CAD tools [9]. In this paper we report an investigation into the use of an evolutionary algorithm (EA) to design genetic circuits which are individually inherently stochastic, but whose behaviour can be predicted at a global level.

Manuscript received January 31, 2010. This work was supported in part by School of Computing Science, Newcastle University, Newcastle upon Tyne UK NE1 7RU (+44 191 222 3864; email j.s.hallinan@ncl.ac.uk; goksel.misirli@ncl.ac.uk; anil.wipat@ncl.ac.uk).

### A. Stochastic systems

Although there is a growing body of literature on stochasticity in genetic systems, the term 'stochastic system' is rarely defined explicitly.

A number of researchers have implicitly defined a stochastic system as one which exhibits different behaviors with different parameter settings. This definition is particularly frequent if the system behavior is critically affected by only one or two parameters [10 - 13], or by slight modification in topology [14]. This definition of stochasticity was used by Voigt *et al.* [15] working with the *Bacillus subtilis sin* operon. These researchers used a deterministic mathematical model and stochastic simulations to investigate the way in which the regulation of the *sin* operon can control population heterogeneity. They showed that different values of one critical parameter led to different dynamic behaviors, ranging from bistable switches through to oscillations, and that "subtle changes in kinetic parameters and the wiring of the network can bias the behavior towards different functions".

Other authors define stochasticity as an evolutionary phenomenon. Under this view mutations produce subtle changes in the wiring pattern and kinetic parameters of a network over the course of several generations. These changes allow the population as a whole to tune its response to a particular environment over generations. Examples include varying the proportion of a population of phage lambda which choose lysis over lysogeny [16], or the proportion of a population of *Bacillus* choosing a particular cell fate [17]. However, we consider this to be evolvability, rather than individual stochasticity, a distinct concept [18]. Both issues are important, but should be considered separately when designing synthetic genetic circuits.

Wilkinson [19] has defined a stochastic computational model as:

> *"A model that contains an element of unpredictability or randomness specified in a precise mathematical way. Each run of a given model will produce different results* [our emphasis], *but the statistical properties of the results of many such runs are pre-determined by the mathematical formulation of the model".*

This definition can usefully be applied to *in vivo* biological systems. We consider a stochastic system to be one which, from a single starting state, can reach multiple stable end states without modification of the system's topology or kinetic parameters.

From the viewpoint of nonlinear dynamics such a system occupies a state space with multiple attractors, as do most

complex systems [20]. However, by our definition a stochastic system is one which is poised close to the confluence of several basins of attraction, and can be diverted into any of them by the action of inherent biological noise. The positioning of the system in state space is determined by its topology and initial parameterization, and slightly different sets of parameter values may affect the tendency of the system to enter different basins of attraction. One set of parameters might lead to an approximately equal probability of each of two possible outcomes for a given system, for example, while a different parameterization might bias the system towards one of the options. As with Wilkinson's stochastic model, the behavior of the population can be predicted with some confidence, while the fate of any particular cell cannot.

A concrete biological example of such a system is cell fate determination in the Gram positive bacterium *Bacillus subtilis*.

### B. *The Bacillus subtilis sin operon*

*B. subtilis* is a ubiquitous, soil-dwelling, non-pathogenic Gram-positive bacterium. Because of its amenability to laboratory manipulation, *B. subtilis* has become the model Gram-positive organism, and its biology and genetics have been extensively studied [21]. Its genome was fully sequenced in 1997 [22] and re-sequenced in 2009 [23].

*B. subtilis* is of particular interest to synthetic biologists because it can exist in a number of different phenotypic states. Cells may produce flagellae and swim as individuals in liquids, swarm in small groups on solids, form multicellular communities composed of bundled chains of sessile cells, form architecturally complex communities embedded in an extracellular matrix (biofilms), or form spores which are inert, resistant to heat and desiccation, and may persist for thousands of years [24]. Under an identical set of environmental conditions, genetically identical cells within a single population may choose different phenotypes; the determination of cell fate is inherently stochastic, making this system ideal for investigations into the engineering of biological stochasticity.

The *sin* operon was identified as the primary controller of sporulation in 1986 [25]. It sits at the heart of a complex web of interactions controlling cell fate decision (Figure 1). The *sin* operon consists of three promoters and two coding sequences. Promoter $P_1$ produces mRNA which codes for two proteins, SinI and SinR. $P_1$ is induced by phosphorylated Spo0A, a two-component response regulator. Spo0A is activated by phosphorylation as part of the general stress response, and is central to the initiation of sporulation. $P_2$ is active after the cell fate decision is made. $P_3$ is weakly constitutive, producing mRNA for SinR. Sin R forms tetramers, which inhibit $P_1$, introducing a negative feedback loop into the system. However, SinR also forms a complex with SinI, reducing its availability to form tetramers.

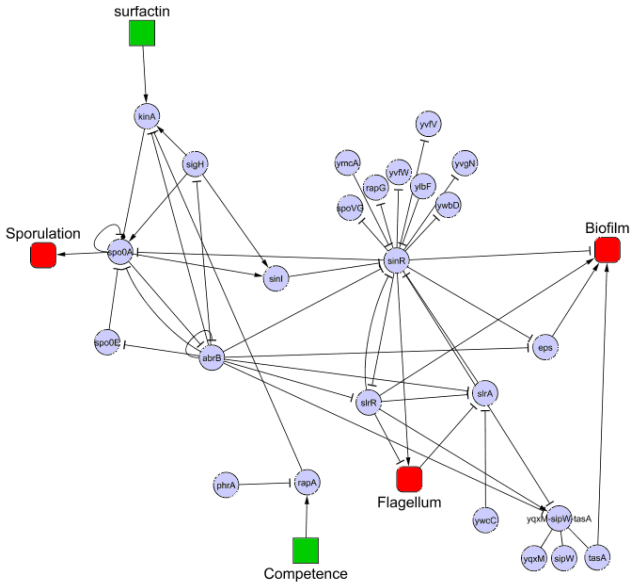


Figure 1. Part of the genetic network controlling cell fate in *Bacillus subtilis*. Possible phenotypes are shown in rounded square boxes; inputs are shown as squares.

SinR is a transcriptional regulator which, amongst other functions, inhibits the transcription of two operons crucial for biofilm formation, *eps* and *yqxM-sipW-tasA*. The *eps* operon produces the extracellular polysaccharides of the biofilm, while one product of the *yqxM-sipW-tasA* operon is TasA. SinR also negatively regulates the production of Spo0A, a major gene in sporulation [26]) The SinR tetramer lies at the heart of a web of inhibitory interactions, and is key to the entire cell fate decision [27].

The amount of SinR tetramer in a cell depends upon the balance between a range of parameters: promoter binding affinities, mRNA and protein degradation rates, complex formation and dissociation rates, and so forth. Different combinations of these parameter values have been demonstrated to lead to different cell fate decisions in a number of models [15, 28, 29], and there is considerable experimental evidence to indicate that they do so in the biological system [30, 31]. The *sin* operon is therefore an excellent target for computational design for synthetic systems. Parameter values measured *in vivo* are available for some parts of the system, but determining which combinations of parameter values lead to which distributions of phenotypes in a population is a non-trivial task.

### C. *Parameterization*

Individual parameter values can fundamentally affect the behaviour of a genetic circuit, the effects of kinetic parameters interact nonlinearly, and multiple sets of parameter values can sometimes produce identical behaviour [12, 32, 33]. Furthermore, the value of a single parameter can differ in the same system under different conditions. Altered kinetics can dramatically affect the behaviour of the system [34]. The same behaviour can be produced by several network topologies, whilst the same topology can produce different behaviours with relatively minor changes in

parameterization [14, 15].

Parameterization is therefore clearly of considerable importance to designers of genetic circuits. Without accurate models of the various component parts of a circuit (promoters, ribosome binding sites, etc) and the way in which they interact (binding affinities, decay rates, etc.) an accurate, predictive model of the entire circuit is impossible.

There are a number of ways in which the issue of model parameterization can be addressed. Measuring parameter values is difficult, for a number of reasons. It is a considerable technical challenge to measure many processes *in vivo*, and measurements made *in vitro* may not be the same as in the living system. Intracellular noise and fluctuations due to the small number of molecules participating in some reactions may also mean that measurements taken at a multi-cellular scale might not reflect conditions in individual cells [35].

For any given model, not all parameter values, or even approximations thereof, will be available from the literature. Missing parameters can be estimated, but since parameterization can profoundly affect the behaviour of a genetic circuit, estimation alone is unlikely to provide biologically plausible results. We collected our initial model parameter values from the literature, and estimated values for missing parameters, and then applied an evolutionary algorithm to fine-tune the parameter values. Since EAs are inherently stochastic each run produces a different set of parameter values producing equivalent behaviour, facilitating sensitivity analysis of the evolved circuit.

### D. *Evolutionary computation for the design of synthetic genetic circuits*

Evolutionary algorithms (EAs) have been extensively used for parameterizing models in a range of domains, including genetic circuit design. EAs have been used for the design of circuit topology [32, 36 - 38], the estimation of parameter values [39, 40], or both [41, 42].

EAs can be used to design the topology of a genetic circuit, to choose parameter values, or both. Circuit topology design is valuable when entirely novel functionality is desired. Parameter value selection is particularly useful in circumstances where the topology of the circuit is known (such as the *sin* operon) but the parameter values are not.

In this project we used an EA to find a set of parameter values for the SinI / SinR system which positions it in such a way that several cell fates are accessible from a single starting state.

## II. METHODS

### A. *The SinI / SinR model*

The model was based upon that developed by Voigt, Wolf & Arkin [15], extended to incorporate the recently-recognised role of SlrR and SlrA, homologues of SinR and SinI respectively [43] (Figure 2). SlrR/SlrA represses $\sigma^D$-dependent flagellar genes and activates the *eps* and *yqxM* operons. The effects of the SinR tetramer and the SlrR protein on the biofilm and flagellar genes was abstracted out, and these relationships replaced with edges representing the assumption that sufficiently high levels of the SinR tetramer inhibit biofilm formation and encourage flagellar growth, whilst the opposite is true for SlrR.

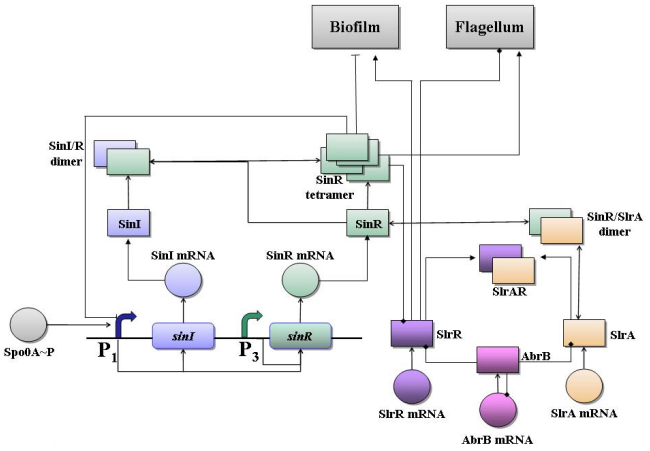


Figure 2. The SinI / SinR model

The model was constructed in the Systems Biology Markup Language (SBML) [44] and simulated using the COPASI biochemical network simulator Version 4.5 build 30 [45], called from a Java EA program via the COPASI Java bindings API.

The model has 53 kinetic parameters, which were initialized using values obtained from the literature. Because the aim of the modelling is to design a physically implementable circuit, only those parameters which can be readily altered *in vivo* were allowed to vary (Table 1).

TABLE 1. KINETIC PARAMETERS VARIED IN THE COURSE OF EVOLUTION

| No. | Description | Initial Value |
|---|---|---|
| 4 | Transcription rate from $P_3$ | 0.02 |
| 6 | Translation rate of SinR | 0.014 |
| 9 | Association rate for SinR tetramers | 0.00125 |
| 15 | Transcription rate of AbrB | 0.2 |
| 17 | Translation rate of AbrB | 0.1 |
| 22 | Transcription rate of SlrA | 0.2 |
| 24 | Translation rate of SlrA | 0.1 |
| 27 | Transcription rate of SlrR | 0.2 |
| 29 | Translation rate of SlrR | 0.1 |
| 34 | Association rate for SlrA:SlrR | 0.0010 |
| 35 | Dissociation rate for SlrA:SlrR | 0.1 |
| 38 | Dissociation rate for SlrA:SinR | 0.1 |
| 47 | Transcription rate of Biofilm gene | 0.2 |

### B. *Evolutionary Algorithm*

The evolutionary algorithm used was a simple Evolutionary Strategy with a population size of two. The problem representation was a string of thirteen real numbers, representing parameter values. This chromosome was initialized to the median of the range of values identified from the literature (Table 1). Each parameterized network

was run 100 times, with a different, small amount of noise, drawn from a normal distribution with mean 0 and variance 0.01, added to the parameter values each time to simulate the effects of biological stochasticity. After 500 timesteps most networks had settled to a stable state, and at this time the final phenotype was identified.

The fitness function depends upon both the proportion of runs which produced each of the two outcomes, and the difference in concentration of the biofilm production indicator protein, *b*, and the flagellum production indicator protein, *f*.

Initially the aim was to produce networks which chose biofilm formation 50% of the time and flagellum formation 50% of the time. The fitness function was therefore based upon the Gaussian distribution:

$$f(x) = ae^{-\frac{(x-p)^2}{2c^2}}, \qquad (1)$$

where *p* is the proportion of runs of a particular network resulting in flagellum development, *a* is the height of the Gaussian, *b* is the position of the centre of the peak, *c* the width of the curve, and *e* is Euler's constant. Since *b = 1.0 – f*, it was not necessary to include *b* in the equation. We set these parameters to produce a Gaussian with maximum value of 1.0 at input *p*, the target proportion of runs producing flagella, for inputs ranging from 0.0 to 1.0 (*a* = 1.0; *b* = *p*; *c* = 0.15). By adjusting *p* we could change the target of the EA.

In order to maximize the difference in expression levels between genes *b* and *f*, the output of the Gaussian was modified by multiplying by the proportional size of *b – f:*

$$fitness = f(x) * \frac{f-b}{f+b} \qquad (2)$$

New networks were then mutated by adding to each parameter an amount drawn from a normal distribution with mean 0.0 and variance 0.1, and the fitness of the child network evaluated as above. Child networks with a fitness equal to or greater than that of the parent replaced the parent. Evolution was continued for 500 generations.

## III. Results

### A. Effect of adding to noise to the simulation

When run using a strictly deterministic algorithm, the SinI / SinR model, with its original parameterization, always triggers the biofilm formation indicator gene (Figure 3).

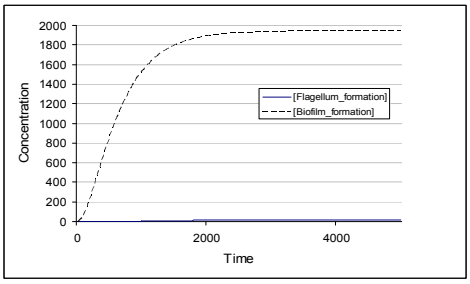


Figure 3. Timecourse of the SinI / SinR model run using a deterministic simulator

However, different runs of the same network under the stochastic evolutionary algorithm produced very different behaviours (Figure 4).

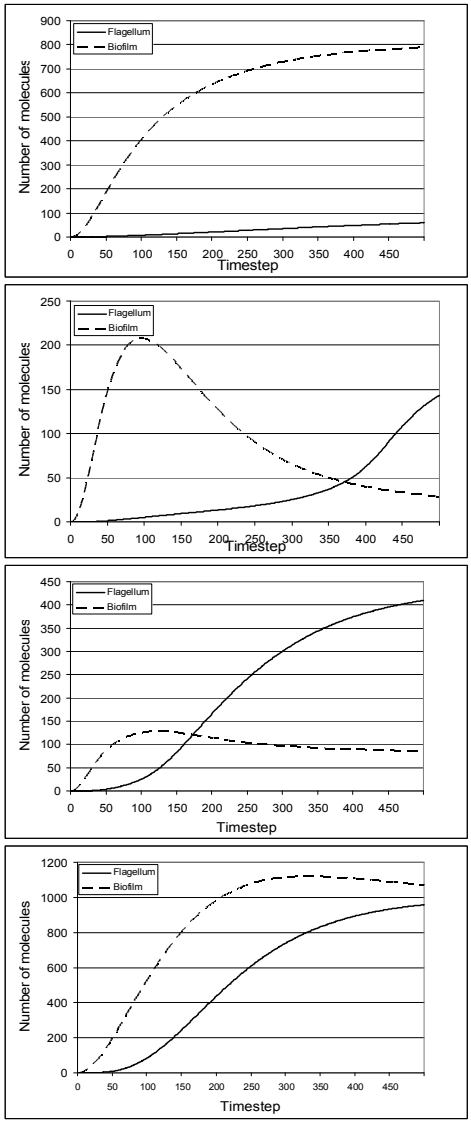


Figure 4. Different runs of the same network with slightly different values for the variable parameters. a) Run 0; b) Run 3; c) Run 5; d) Run 7. Note the different scales on the y axes.

### B. Evolving to different targets

We ran the systems with three different target outcomes: a proportion of runs producing flagella of 0.8 (80F); 0.5 (50F); or 0.2 (20F). The course of evolution for the three targets was quite different (Figure 5).

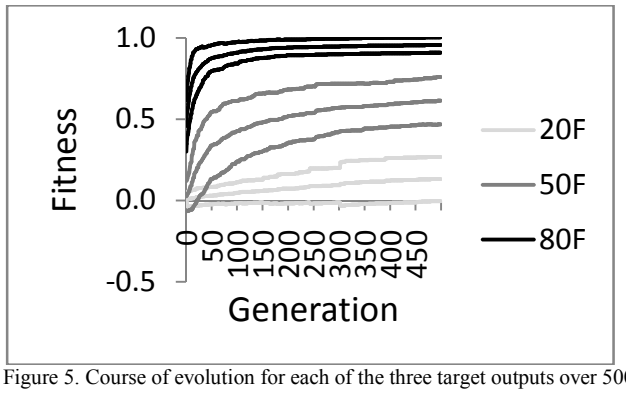

Figure 5. Course of evolution for each of the three target outputs over 500 generations

Each network was run 100 times, for 500 steps each time. The distribution of fitness values achieved at the end of the runs was also very different for the different targets (Figure 6).

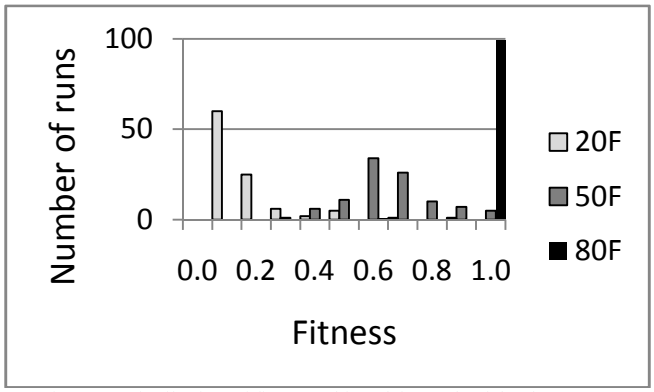

Figure 6. Distribution of fitness achieved for each of the three target outcomes.

The proportion of runs producing flagella also differed for each of the three targets (Figure 7).

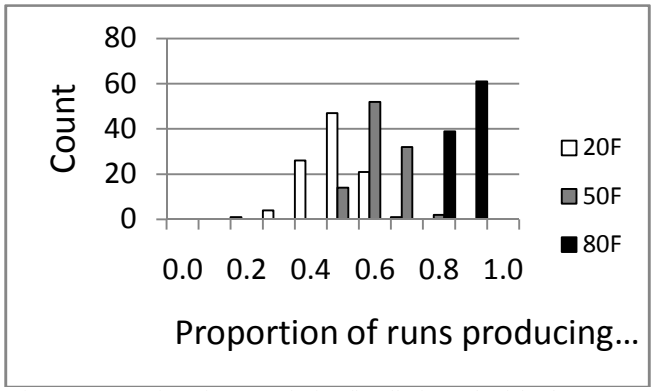

Figure 7. Proportion of runs producing flagella for each of the three targets

### C. Changes in parameter values

Considerable variability was observed in the final parameter values in the evolved models (Figure 8); the association rate for SinR tetramers (P9) varied most in evolving the model towards meeting the species ratio of biofilm producing to motile cells. Increasing the affinity of the SinR monomers for each other in tetramer formation increased the probability of single cell motility, whereas increasing the affinity of SlrA for SlrR resulted in a higher probability of biofilm formation. Significant variations in the strength of the $P_3$ promoter (P4) and the translation rate of SinR (P6) were also noted during the evolution of the model, but with little obvious selective behaviour in terms of the final behaviour of the evolved model.

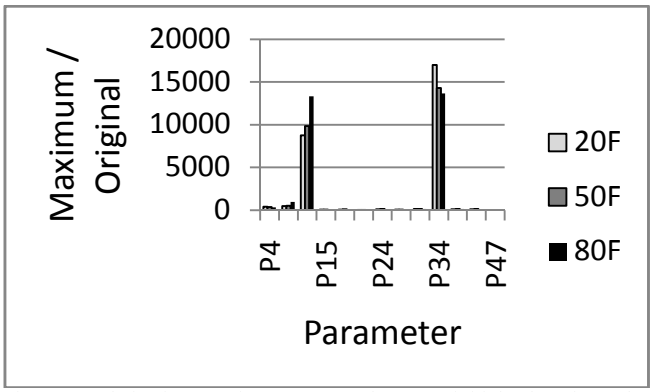

Figure 8. Relative increase in parameter values over initial values for each of the three target outputs. P9 is the association rate for SinR tetramers, while P34 is the association rate for SlrA:SlrR

## IV. Discussion

Computational evolution for the design of a biological circuit such as the *sin* operon and its local interactome has the potential to provide information about both the characteristics of the circuit, and the way in which it functions. Our model suggests that the *sin* operon switch is biased: it can be evolved very easily to produce a large proportion of flagellated cells; less easily to produce an intermediate proportion; and only with difficulty to produce a large proportion of cells choosing the biofilm fate (Figure 7). This bias appears to make evolutionary sense, since the production of a flagellum, and subsequent motility and escape from difficult local conditions is likely to be less of an energy commitment for the cell than is biofilm formation. Further, not all of the cells involved in biofilm formation actually choose the biofilm phenotype; in fact, only a minority do so, and their contribution is sufficient to maintain the biofilm for the entire community [46].

However, the deterministic model produces exactly the opposite result; with the parameterisation we initially used the model moves smoothly and swiftly to biofilm formation. This behaviour could merely be due to a non-optimal choice of parameter values, always a major issue when setting up models of circuits whose behaviour has not been studied in depth. However, the fact that the deterministic model always selects one phenotype, whilst real populations of cells choose between any of a number of fates, indicates that this is not just a parameter value problem. Stochastic processes are important in real cells, and some measure of stochasticity is clearly crucial to the model achieving biologically plausible behaviour.

The "stochastic" algorithm we used is not truly stochastic. By combining a deterministic simulation algorithm with a

slight change of parameter values, sampled from a Gaussian distribution, for each run of a given network, we aimed to construct a system which was computationally efficient enough to allow large numbers of runs to be performed and analyzed, while still retaining stochastic characteristics. This strategy lead to the observation of the biased nature of the switch, a finding which we are currently investigating *in vivo*.

The mechanism of operation of the switch is also of considerable interest. It is clear that feedback loops and inhibition are critical to the behaviour of the system, an observation which is not unexpected, since the presence of feedback loops is known to be important to both multistability and oscillatory dynamics in genetic circuits [47, 48]. However, the fact that control in this circuit is primarily via complex formation is not apparent in the conventional, network-centric view of Figure 1.

The SinR tetramer is the active form of the protein, and it is "inhibited" not primarily by transcriptional control, but by the formation of complexes with SinI and SlrR. Such a mechanism introduces a wide range of factors into the control process. Parameters such as transcription, translation degradation and association rates of the various molecular species involved can all be modified individually, potentially providing much more fine-grained and flexible control over the behavior of the switch than is possible simply by transcriptional regulation. Our results indicate that complex association rates are particularly important for this particular switch (Figure 8).

The *sin* operon switch is therefore a complex system poised at the confluence of multiple basins of attraction in phenotype space. Manipulation of its parameters is sufficient to alter the extent to which it will choose one phenotype over another under identical conditions, making it a potentially valuable tool for the synthetic biologist.

There are a number of applications for a tunable, stochastic switch in synthetic biology. Although bacteria are often considered as independent, isolated individuals, they in fact exist embedded in a web of interactions with conspecifics, microbes from other genera, and even organisms from other kingdoms. A bacterial community can thus function almost as a multicellular organism, and under these conditions the ability to engineer the behavior of the population as a whole becomes of fundamental importance. The 2009 Newcastle iGEM team[1], for example, aimed to engineer *B. subtilis* to take up cadmium from its environment, sequester it into spores, and disable germination of the spores. In order to maintain a balance between spore-forming (and therefore inviable) cells and vegetatively growing cells, it was important to include a tuneable stochastic switch into the design.

In this work we concentrated upon an analysis of the parameters which could be modified *in vivo*. Another possibility for modification of the switch is to engineer changes to the network topology of the model. With the advent of large-scale, reasonably-priced DNA synthesis technology, relatively large changes to a system can be made upon the basis of computational design. For example, interactions between a transcription factor and a gene can be added or eliminated by changing the promoter of the target gene, while mRNA and protein levels can be modified by changing the binding affinities of the promoter or ribosome binding site. We are currently investigating extending our model to include wiring changes, and an analysis of the interaction between circuit topology and parameterization. We will also extend the model to include more than the two possible phenotypes—flagellum or biofilm—currently supported.

We have investigated the feasibility of engineering the *sin* operon switch controlling cell fate in *B. subtilis* to produce a predictable proportion of the population choosing a particular phenotype. We find that the switch appears to be naturally biased towards flagellum formation, an observation which tallies well with experimental data. It is possible to modify the switch such that this biased is changed, and even reversed, although many of the network runs did not achieve this task. Complex association rates appear to be of primary importance in engineering the switch, confirming the primacy of complex formation over transcriptional regulation in the function of the *sin* stochastic switch.

[1] http://2009.igem.org/Team:Newcastle

## REFERENCES


[1] Vilar, J. M. G., Kueh, H. Y., Barkai, N., and Leibler, S. "Mechanisms of noise-resistance in genetic oscillators", *Proceedings of the National Academy of Sciences of the USA* vol. 99, no. 9, pp.5988 - 5992, 2002.

[2] Bollenbach, T. & Kishony, R. (2009). Quiet gene circuit more fragile than its noisy peer. *Cell* 139: 460 - 461.

[3] Maheshri, N., and O"Shea, E. K. "Living with noisy genes: How cells function reliably with inherent variability in gene expression", *Annual Review of Biophysics and Biomolecular Structure* vol. 36, pp.413 - 434, 2007

[4] Cagatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J., and Suel, G. M. "Architecture-dependent noise discriminates functionally analogous differentiation circuits", *Cell* vol. 139, pp.512 - 522, 2009.

[5] Endy, D. "Foundations for engineering biology", *Nature* vol. 438, pp.449 - 453, 2005.

[6] Glass, J. I., Assad-Garcia, N., Alperovitch, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchinson, C. A., Smith, H. O. & Venter, J. C. "Essential genes of a minimal bacterium", *Proceedings of the National Academy of Sciences of the USA* vol. 103, no. 2, pp.425 - 434, 2006.

[7] Lartigue, C., Vashee, S., Algire, M. A., Chuang, R.-Y., Benders, G. A. & Ma, L. "Creating bacterial strains from genomes that have been grown and engineered in yeast", *Science* vol. 325, no. 5948, pp.1693 - 1696, 2009.

[8] Pumick, P. E. M., and Weiss, R. "The second wave of synthetic biology: From modules to systems.", *Nature Reviews Molecular Cell Biology* vol. 10, pp.410 - 422, 2009

[9] Heinemann, M., and Panke, S. "Synthetic biology - putting engineering into biology", *Bioinformatics* vol. 22, no. 22, pp.2790 - 2799, 2006.

[10] Ozbudak, E. M., Thattai, M., Lim, H. N., and van Oudenaarden, A. "Multistability in the lactose utilization network of *Escherichia coli*", *Nature* vol. 427, pp.737 - 740, 2004.

[11] Setty, Y. E., Mayo, A., Surette, M. G., and Alon, U. "Detailed map of a cis-regulatory input function", *Proceedings of the National Academy of Sciences of the USA* vol. 100, pp.7702 - 7707, 2003.

[12] von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M. "The segment polarity network is a robust developmental module", *Nature* vol. 406, pp.188 - 192, 2000.

[13] Meir, E., von Dassau, G., Munro, E., and O'Dell, G. M. "Robustness, flexibility and the role of lateral inhibition in the neurogenic network", *Current Biology* vol. 12, pp.778 - 785, 2002.
[14] Atkinson, M. R., Savageau, M. A., Myers, J. T., and Ninfa, A. J. "Development of genetic circuitry exhibiting toggle switch or oscillatory behaviour in *Escherichia coli", Cell* vol. 113, pp.597 - 607, 2003.
[15] Voigt, C. A., Wolf, D. M., and Arkin, A. P. "The *Bacillus subtilis sin* operon: An evolveable network motif", *Genetics* vol. 169, pp.1187 - 1202, 2005.
[16] Little, J. W., Shepley, D. P., and Wert, D. W. "Robustness of a gene regulatory circuit", *EMBO Journal* vol. 18, pp.4299 - 4307, 1999.
[17] Maughan, H., and Nicholson, W. L. "Stochastic processes influence stationary-phase decisions in *Bacillus subtilis", Journal of Bacteriology* vol. 186, pp.2212 - 2214, 2004.
[18] Bedau, M. A., and Packard, N. H. "Evolution of evolvability via adaptation of mutation rates", *Biosystems* vol. 69, pp.143 - 162, 2003.
[19] Wilkinson, D. J. "Stochastic modelling for quantitative description of heterogeneous biological systems", *Nature Reviews Genetics* vol. 10, pp.122 - 133, 2009.
[20] Nusse, H. E., and Yorke, J. E. "Basins of attraction", *Science* vol. 5254, pp.1376 - 1380, 1996.
[21] Anagnostopoulos, C., and Spizizen, J. "Requirements for transformation in *Bacillus subtilis", Journal of Bacteriology* vol. 81, no. 5, pp.741 - 746, 1961.
[22] Kunst, F., *et al.* "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis", Nature* vol. 390, no. 6657, pp.234 - 256, 1997.
[23] Barbe, V., Cruveiller, S. Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T. Moszer, I., Médigue, C. & Danchin, A. "From a consortium sequence to a unified sequence: The Bacillus subtilis 168 reference genome a decade later", *Microbiology* vol. 155, pp.1758 - 1775, 2009.
[24] Sonenshein, A. L., Hoch, J. & Losick, R. (eds.) *"Bacillus subtilis and Its Closest Relatives: From Genes to Cells"* : American Society for Microbiology, 2002.
[25] Gaur, N. K., Dubnau, E., and Smith, I. "Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies", *Journal of Bacteriology* vol. 168, no. 2, pp.860 - 869, 1986.
[26] Lewis, R. J., Brannigan, J. A., Offen, W. A., Smith, I., and Wilkinson, A. J. "An evolutionary link between sporulation and prophage induction in the structure of a repressor: anti-repressor complex", *Journal of Molecular Biology* vol. 283, pp.907 - 912, 1998.
[27] Errington, J. *"Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis", *Microbiological Reviews* vol. 57, no. 1, pp.1 - 33, 1993.
[28] de Jong, H., Geiselman, J., Batt, G., Hernandez, C., and Page, M. "Qualitative simulation of the initiation of sporulation in *Bacillus subtilis", Bulletin of Mathematical Biology* vol. 66, pp.261 - 299, 2004
[29] Morohashi, M., Ohashi, Y., Tani, S., Ishii, K., Itaya, M., Nanamiya, H., Kawamura, F., Tomita, M. & Soga, T. "Model-based definition of population heterogeneity and its effects on metabolism in sporulating *Bacillus subtilis", Journal of Biochemistry* vol. 142, no. 2, pp.183 - 191, 2007.
[30] Iber, D., Clarkson, J., Yudkin, M. D., and Campbell, I. D. "The mechanism of cell differentiation in *Bacillus subtilis", Nature* vol. 441, pp.371 - 374, 2006.
[31] Losick, R., and Desplan, C. "Stochasticity and cell fate", *Science* vol. 320, pp.65 - 68, 2008.
[32] Tomshine, J., and Kazneiss, Y. N.. "Optimization of a stochastically simulated gene network model via simulated annealing", *Biophysical Journal* vol. 91, pp.3196 - 3205, 2006.
[33] Elowitz, M. B., and Leibler, S. "A synthetic oscillatory network of transcriptional regulators", *Nature* vol. 403, pp.335 - 338, 2000.
[34] Guet, C. C., Elowitz, M. B., Hsing, W., and Leibler, S. "Combinatorial synthesis of gene networks", *Science* vol. 296, no. 5572, pp.1466 - 1470, 2002.
[35] Hasty, J., Dolnik, M., Rottschafer, V., and Collins, J. J. "Synthetic gene networks for entraining and amplifying cellular oscillations", *Physical Review Letters* vol. 88, no. 14, pp.148101-1 - 148101-4, 2002.
[36] Deckard, A., and Sauro, H. M. "Preliminary studies on the *in silico* evolution of biochemical networks", *ChemBioChem* vol. 5, pp.1423 - 1431, 2004.
[37] Paladuqu, S.R., Chicarmane, V., Deckard, A., Frumkin, J. P., McCormack, M. & Sauro, H. M. "In silico evolution of functional modules in biochemical networks", *Systems Biology (Stevenage)* vol. 153, no. 4, pp.223 - 235, 2006.
[38] Chu, D. "Evolving genetic regulatory networks for systems biology", *"Proceedings of the 2007 Congress on Evolutionary Computation (CEC 2007)"* : pp.875 - 882, 2007.
[39] Gilman, A., and Ross, J. "Genetic-algorithm selection of a regulatory structure that directs flux in a simple metabolic model", *Biophysical Journal* vol. 69, no. 4, pp.1321, 1995.
[40] Tsuchiya, M., and Ross, J. "Application of genetic algorithm to chemical kinetics: Systematic determination of reaction mechanism and rate coefficients for a complex reaction network", *Journal of Physical Chemistry A* vol. 105, no. 16, pp.4052 - 4058, 2001.
[41] Yokobayashi, Y., Collins, C. H., Leadbetter, J. R., Weiss, R., and Arnold, F. H. "Evolutionary design of genetic circuits and cell-cell communications.", *Advances in Complex Systems* vol. 6, no. 1, pp.1 - 9, 2003.
[42] Lenser, T., Hinze, T., Ibrahim, B., and Dittrich, P. "Towards evolutionary network reconstruction tools for systems biology", In Marchiori, E., Moore, J. H. & Rajapakse, J. C. (eds.) *"EvoBio 2007, Lecture Notes in Computer Science* 4447*"* : pp.132 - 142, 2007.
[43] Kobayashi, K. "SlrR / SlrA controls the initiation of biofilm formation in *Bacillus subtilis", Molecular Microbiology* vol. 69, no. 6, pp.1399 - 1410, 2008.
[44] Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C. et al."The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models", *Bioinformatics* vol. 19, pp.524 - 531, 2003.
[45] Hoops, S., Sahle, S., Guages, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P. & Kummer, U. "COPASI-a COmplex PAthway SImulator", *Bioinformatics* vol. 22, no. 24, pp.3067 - 3074, 2006.
[46] Chai, Y., Chu, F., Kolter, R., and Losick, R. "Bistability and biofilm formation in *Bacillus subtilis", Molecular Microbiology* vol. 67, no. 2, pp.254 - 263, 2008.
[47] Angeli, D., Ferrell, J. E., and Sontag, E. D. "Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems", *Proceedings of the National Academy of Sciences of the USA* vol. 101, no. 7, pp.1822 - 1827, 2004.
[48] Pigolotti, S., Krishna, S., and Jensen, M. H. "Oscillation patterns in negative feedback loops", *Proceedings of the National Academy of Sciences of the USA* vol. 104, no. 16, pp.6533 - 6537, 2007.