Engineered gene circuits

Jeff Hasty*, David McMillen† & J. J. Collins†

**Department of Bioengineering, University of California San Diego, La Jolla, California 92093, USA (e-mail: hasty@ucsd.edu)*
†Center for BioDynamics and Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA

A central focus of postgenomic research will be to understand how cellular phenomena arise from the connectivity of genes and proteins. This connectivity generates molecular network diagrams that resemble complex electrical circuits, and a systematic understanding will require the development of a mathematical framework for describing the circuitry. From an engineering perspective, the natural path towards such a framework is the construction and analysis of the underlying submodules that constitute the network. Recent experimental advances in both sequencing and genetic engineering have made this approach feasible through the design and implementation of synthetic gene networks amenable to mathematical modelling and quantitative analysis. These developments have signalled the emergence of a gene circuit discipline, which provides a framework for predicting and evaluating the dynamics of cellular processes. Synthetic gene networks will also lead to new logical forms of cellular control, which could have important applications in functional genomics, nanotechnology, and gene and cell therapy.

It has been over 40 years since Monod and Jacob boldly predicted that such fundamental cellular processes as differentiation and protein regulation are accomplished through signalling pathways resident at the level of the gene[1]. This prediction laid the foundation for the ensuing progress in describing the essential regulatory mechanisms in many specific genetic systems. With the development of the field of nonlinear dynamics and the concurrent advent of significant computing power, mathematical models describing gene regulation began to appear regularly in the 1970s[2–9]. Implicit in these studies was the realization that the 'wiring' of naturally occurring gene regulatory networks would be too complex for qualitative descriptions devoid of mathematics. Although this realization proved to be ahead of its time, owing mainly to the lack of experimentally deduced regulatory pathways in the 'pre-genomic' era, recent experimental advances have re-ignited interest in the development of circuit analysis techniques for describing complex gene networks.

The concept of designed gene circuits has motivated researchers to draw direct analogies with established techniques in electrical engineering[10,11]. As with the construction of electrical circuits, the gene circuit approach[12–24] uses mathematical and computational tools in the analysis of a proposed circuit diagram, while novel experimental techniques are used to construct the networks according to the model blueprint. So far, the qualitative agreement between model and experiment in a series of studies[13–15,22,23] has supported the notion of such an engineering-based methodology (for a detailed discussion of the various mathematical modelling techniques and their particular applications, see refs 17 and 19). The power of this approach is that it can be used to study simplified systems to gain insight into the general 'modules' of gene regulation[25–27]. These modules include subnetworks that act as switches or oscillators, as well as networks that act to communicate across a population of cells. This ability to engineer gene networks offers the prospect of extracting carefully chosen subsystems from natural organisms, and focusing both modelling and experimental effort on determination of the behaviour of the subsystems in isolation. Furthermore, there is the possibility of using the insights thus obtained to create genetic 'control systems', designed to correct faulty cellular mechanisms, or to generate entirely new modes of behaviour.

Autoregulatory systems

Feedback loops are an important concept in engineering control systems. In the context of gene regulation, feedback occurs through autoregulation, wherein a protein modifies, directly or indirectly, its own rate of production. Whether such interactions embody positive or negative feedback depends on the details of the network dynamics. Understanding the nature of such feedback loops in biological networks is a key step in the attempt to formulate a gene circuit discipline[5–9,28–30].

Seminal work in the modelling of gene networks[3] focused on the stability properties of networks dominated by positive versus negative feedback. Stability refers to the tendency of a system to remain close to a steady state (a state in which production and decay rates are balanced) despite the influence of perturbations. A central result of this work was that genes regulated by negative feedback should be more stable than either unregulated genes or those regulated by positive feedback (an example of a naturally occurring negative feedback system is given in Fig. 1a). In the past few years, synthetic gene networks have been engineered to test the first portion of this prediction, comparing the behaviour of a simplified gene regulatory network based on negative feedback to the behaviour of the equivalent unregulated network[15].

The experiments used a promoter that is shut off by the tetracycline repressor protein (TetR) to control the production of TetR, and compared this network to an unregulated system (Fig. 1b). Using a fusion of green fluorescent protein (GFP) to the TetR protein allowed observation of the state of the network (that is, the number of TetR molecules present in the cell) through fluorescence microscopy. Sampling multiple cells yielded a distribution of fluorescence intensities, and the stability of the steady state was evaluated by the width of these distributions. Because the degree of stability is inversely proportional to the width of the distribution, narrower distributions imply greater stability. Mathematical modelling indicated, as had the earlier analysis, that the negative feedback network should be more stable than the unregulated network, and the experimental results confirmed this prediction.

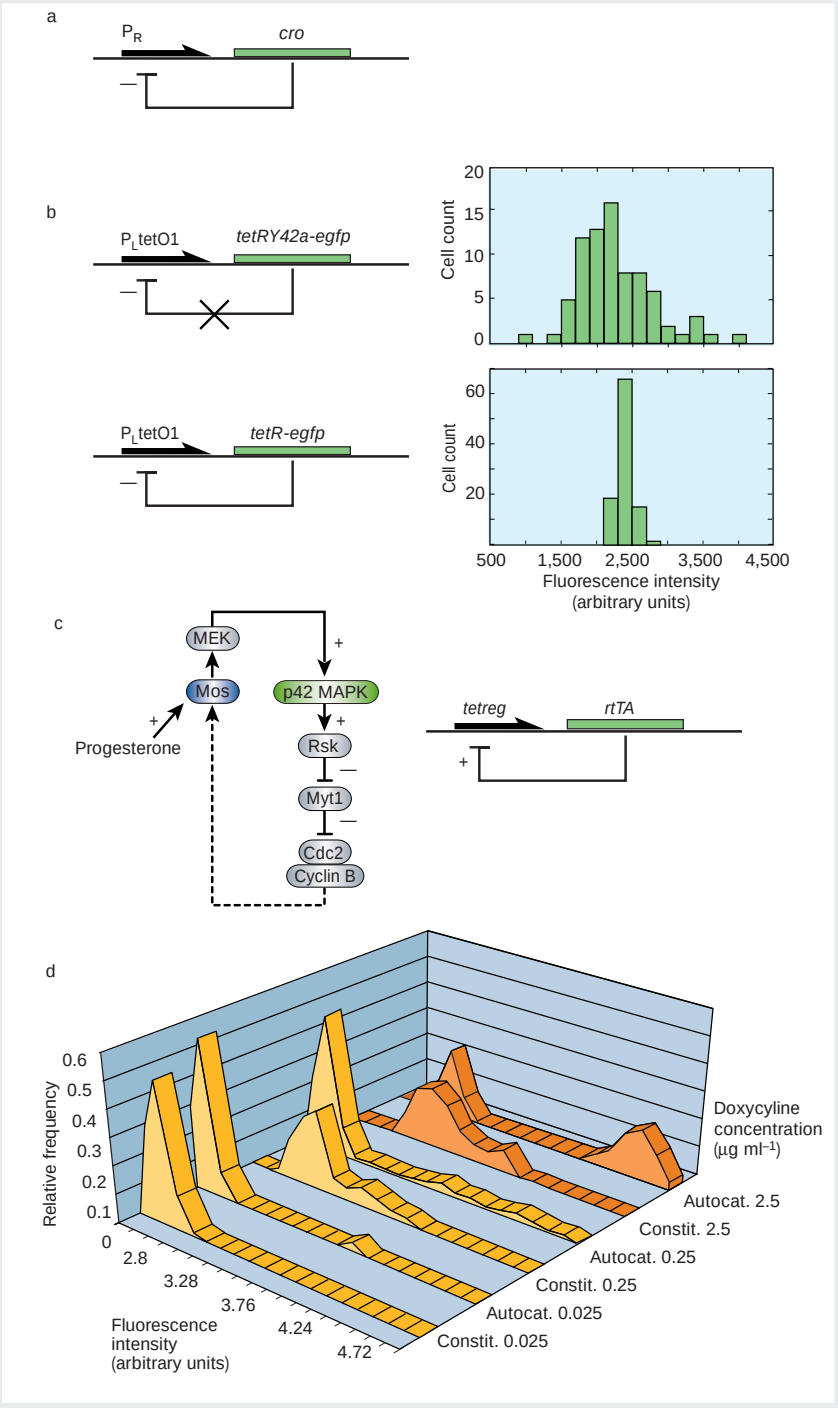

Figure 1 Autoregulatory systems. **a**, In this natural negative feedback system from the bacteriophage λ, the promoter P_R controls the expression of the Cro protein, which represses P_R. **b**, The synthetic negative feedback system[15] uses the promoter $P_L$tetO1 to control the expression of TetR-EGFP, a fusion of the tetracycline repressor (TetR) and the enhanced green fluorescent protein (EGFP). Negative feedback arises because TetR represses transcription from $P_L$tetO1. Replacing TetR with TetRY42A eliminates the feedback, producing an unregulated system. The distribution of observed expression states for the unregulated system (upper, right) is about three times wider than the distribution for the negative feedback system (lower, right), demonstrating improved stability with negative feedback. (Distributions redrawn from ref. 15.) **c** The left panel shows a natural positive feedback system, the Mos-MEK-p42 MAPK cascade, which controls part of the maturation process in *Xenopus* oocytes. Progesterone stimulates the production of the Mos protein, which indirectly activates p42 MAPK (mitogen-activated protein kinase). p42 MAPK activation, in turn, stimulates production of Mos through a series of steps, not fully known (dashed line indicates unknown intermediates). (Redrawn from ref. 31.) The right panel shows a synthetic positive feedback system[22] in which the promoter region *tetreg* controls expression of the tetracycline-responsive transactivator (rtTA); rtTA activates *tetreg*, completing the positive feedback loop. **d**, Observed bistability in the synthetic positive feedback system[22]. Fluorescence intensities are shown for the positive feedback (autocat., autocatalytic) and unregulated (constit., constitutive) systems. The concentration of the inducer doxycycline controls the degree of positive feedback, as regulatory binding of rtTA relies on activation by the inducer. For low inducer concentrations (yellow), both the constitutive and positive feedback systems have distributions with a single peak. At higher concentrations of inducer (orange), the constitutive system remains unimodal, while the strong positive feedback causes the autocatalytic system to split into two distinct populations of cells. (Redrawn from ref. 22.)

A significant feature of positive feedback is its role in the generation of bistability, where two steady states of the system are stable (see ref. 31 for a recent review of the role of positive feedback in bistability of gene regulatory networks). The importance of positive feedback in generating multiple stable states has been analysed mathematically[30,32] and has been implicated in the stability of the differentiated and undifferentiated states in *Xenopus* oocytes[33,34] (Fig. 1c). Experiments with an engineered positive feedback network[22] have demonstrated the existence of bistability in the system (Fig. 1d). The synthetic network was implemented in the budding yeast *Saccharomyces cerevisiae*, and consisted of a tetracycline-responsive transactivator (rtTA) that activated its own promoter. As in the negative feedback experiments, the reporter protein GFP was fused to the transactivator to allow observation of expression levels through fluorescence microscopy. As predicted by the accompanying mathematical model[22], the resulting distributions were bimodal: there were two distinct subpopulations of cells, with one group expressing small amounts of the protein, whereas the other expressed large amounts.

In the synthetic positive feedback network[22], the partitioning of the cells into two subpopulations was not permanent; this was attributed to fluctuations in the network that were large enough to cause spontaneous transitions from one state to the other. Only transitions from the low-expression state to the high-expression state were observed experimentally, but there is no theoretical reason

precluding transitions in the opposite direction, and it may be chance that no such transitions occurred during the period of observation. It is well known that noise can drive a system back and forth between two stable states, and the average time it takes for such a transition to occur is called the 'escape time'. The escape time is a function of both the stability of the states and the size of the fluctuations.

In the case of *Xenopus* oocytes, differentiation is an irreversible process: once the egg cells mature, they are never observed to change back to the immature state. This indicates that the escape time of the maturation system is either infinite, or so long that it is effectively infinite; that is, over the lifetime of the organism, there is a negligible chance of making a spontaneous transition back to the immature state. (An alternative possibility in a complex network of this sort is that once maturation is achieved, the system parameters change in such a way that the bistability is irreversibly eliminated.) The fact that the synthetic positive feedback network[22] made transitions on a significantly shorter timescale suggests that it was either subject to greater noise, or that its expression states were less stable. Controlled experiments on synthetic autoregulatory networks, combined with theoretical treatments[16,18,35,36], may serve to identify the precise differences between switching systems.

Toggle switch

Bistability is a minimal requirement for a network to possess memory, where the state of the network stores information about its past. When forced by a transient stimulus into one state or the other, such a system remains in that state after the transient has been removed, thus 'remembering' the stimulus event. For generating bistability, an alternative to the positive feedback network is mutual inhibition. This method of achieving bistability arises in a number of contexts: in engineering, there is the Reset–Set latch (widely known as the 'RS latch') circuit design, and switches based on mutual repression have long been suggested as a common element in gene regulatory networks[1]. One such genetic switch is found in the P_R/P_{RM} region in λ phage, which acts co-repressively to control the lysis/lysogeny decision: Cro, controlled by P_R, represses P_{RM}, whereas CI, controlled by P_{RM}, represses P_R (Fig. 2a).

The principle of mutual repression was used to achieve bistability in a synthetic genetic toggle switch[13]. Its design made use of a mathematical model to deduce the parameter regimes required for bistability and robust switching. These criteria included the use of strong and balanced constitutive promoters, effective transcriptional repression, the formation of protein multimers, and similar protein degradation rates for the two main components.

An example of one of the toggle switch designs is shown in Fig. 2b. In this version of the toggle, the *lac* gene is under the control of the P_Ls1con promoter, whereas the *cI* gene is controlled by the Ptrc-2 promoter; the lactose repressor (LacI) protein represses Ptrc-2, and the CI protein represses P_Ls1con. Experimentally, switching between the two states was induced by the transient application of either a chemical or thermal stimulus (Fig. 2c). The chemical inducer was isopropyl-β-D-thiogalactopyranoside (IPTG), which binds to LacI tetramers and renders them effectively unable to repress Ptrc-2. A temperature-sensitive version of the CI protein was used, so that protein denaturation increased with temperature, allowing a thermal stimulus to eliminate the active CI in the system.

Logic gates

The concept of engineered gene circuits has led to a formulation based on logic gates and their associated truth tables, with resulting schematics that are the direct analogue of electronic circuit diagrams[37,38]. One such description defines the inputs to a regulated promoter as the protein/inducer pair, and the output as 'on' if the gene downstream of the promoter is being transcribed and 'off' otherwise (Fig. 3a). For example, consider the arabinose operon, which is induced by a complex consisting of AraC dimers and the chemical arabinose. The inputs are AraC proteins and aribinose, and the arabinose promoter is on only if AraC and arabinose are present, and is off otherwise. In the parlance of electrical engineering, the truth table for such a system of inputs and outputs leads to a logical AND gate. Simple logic gates, such as AND gates and OR gates (see Fig. 3b), can be combined to yield circuits of any given complexity, and indeed a central focus of ref. 38 was the formulation of an engineering circuit discipline with a simulation package for analysing the resulting gene circuits[38].

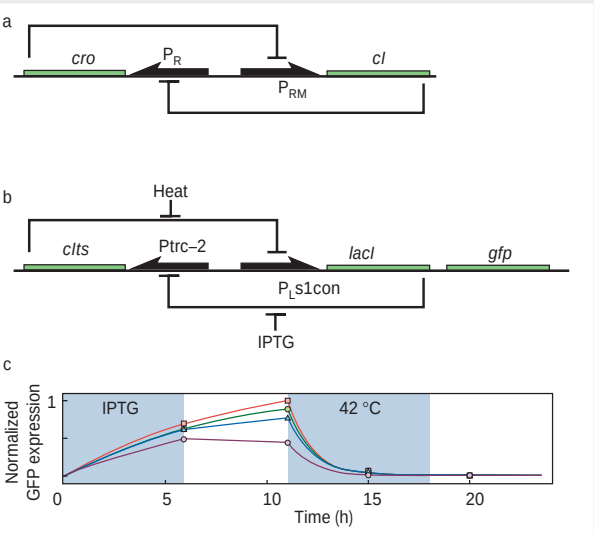

Figure 2 Natural and synthetic co-repressive switches. **a**, In this natural switch from the bacteriophage λ, the promoters P_{RM} and P_R are each repressed by the product of the other: P_{RM} controls expression of the gene *cI*, and the protein CI represses P_R, whereas P_R controls expression of the gene *cro*, and the protein Cro represses P_{RM}. **b**, The synthetic genetic toggle switch[13] uses the promoter Ptrc-2 to control the production of a temperature-sensitive version of the CI protein (expressed by *cIts*); CI acts to repress the promoter P_Ls1con. Conversely, P_Ls1con controls transcription of the gene *lacI*, whose product LacI (lactose repressor) represses Ptrc-2. **c**, Experimental results showing bistability of a genetic toggle switch in *Escherichia coli*[13]. The response of green fluorescent protein (GFP) is shown, which corresponds to expression of the *cI* gene. Shaded regions indicate periods of induced switching. The cells were forced to the high-GFP state by exposure to isopropyl-β-D-thiogalactopyranoside (IPTG), which eliminates the repressive effect of LacI; note that the cells remain in the high state after the inducer is removed. The population was then forced into the low-GFP state by induction with a thermal pulse, which eliminated the active CI in the system by increasing the denaturation rate of the temperature-sensitive CI protein. Upon returning to the base temperature, the cells one again remain in the switched state after the induction is removed. The different coloured lines represent different plasmid strains implementing slight variations of the toggle. (Redrawn from ref. 13.)

Although the description of logic gates in terms of protein-chemical inputs is one possible approach, a complementary formulation involves defining two external chemicals as the input signals. For example, consider the schematic for an alternative AND gate depicted in Fig. 3b (F. J. Isaacs, C. R. Cantor and J.J.C., manuscript in preparation). The circuitry is such that the first promoter directs the polycistronic transcription of the *lac* and *tet* genes, and the second promoter is engineered to be repressed by either LacI or TetR. When the two chemicals IPTG and anhydrotetracycline (aTc) are present, the LacI and TetR repressors, respectively, are inactivated and the *gfp* gene downstream of promoter 2 is transcribed. Thus, the circuit forms an AND gate as both chemical inputs (IPTG and aTc) must be present for promoter 2 to be on (GFP expressed).

A central theme in gene circuit design is that the simpler 'fundamental' circuits form the basis for more complex designs. For example, memory can be added to the AND network by letting the second promoter direct the production of a third protein capable of

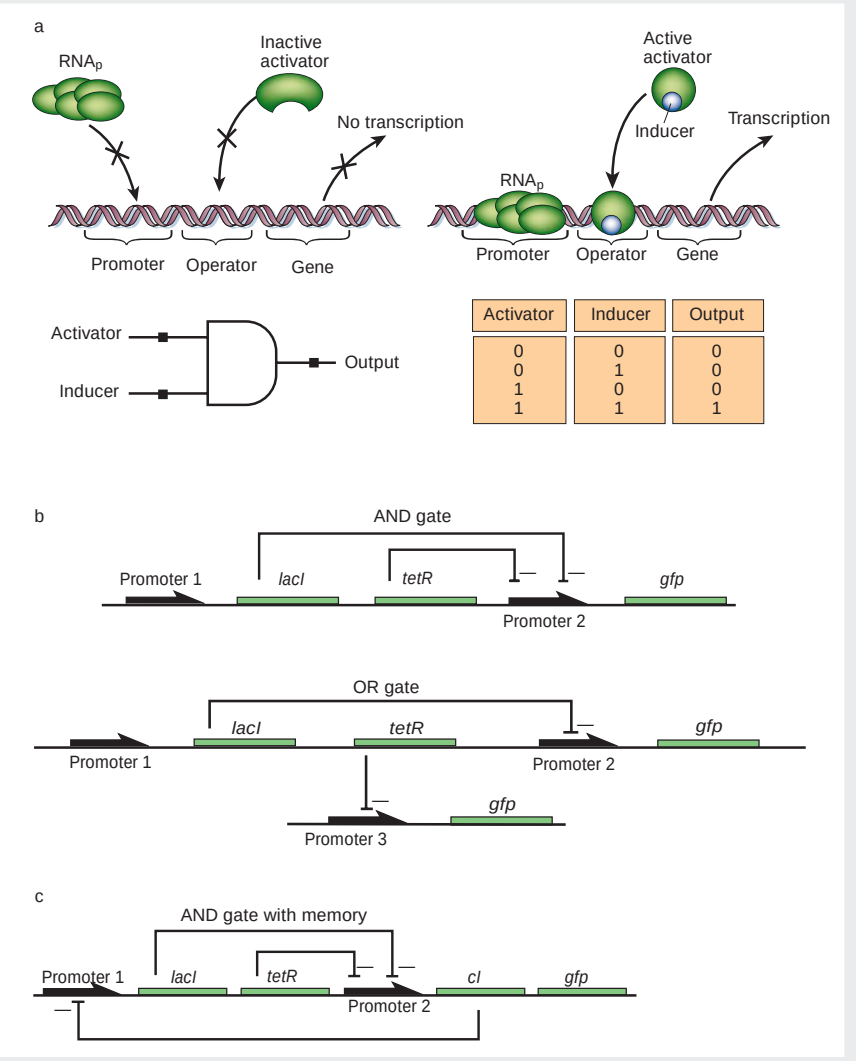

Figure 3 Logic gates. **a**, Genetic and electronic circuit diagrams for an AND gate using proteins and inducers for the inputs, and the state of the gene (on or off) as the output. The corresponding truth table elucidates the logic of the AND gate. **b**, An alternative AND gate uses two inducers as inputs and expression of the green fluorescent protein (*gfp*) gene as the output. The *lacI* and *tetR* genes (encoding tetracycline and lactose repressor, respectively) are expressed polycistronically by a constitutive promoter. If either LacI or TetR bind to the second promoter, the expression of the *gfp* gene is turned off. Because both of the inducers isopropyl-β-D-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc) are needed to prohibit the repression of the second promoter by LacI and TetR, the circuit forms an AND gate. The OR gate similarly uses the polycistronic expression of *lacI* and *tetR* from the first promoter, but differs in that only LacI represses the second promoter, while TetR represses a third promoter. Because the presence of either inducer (or both) leads to the expression of the *gfp* gene, the circuit forms an OR gate. **c**, Memory in the AND gate is achieved by inserting an additional *cI* gene under the control of the second promoter, and using a CI-repressible first promoter. Once the system is switched into the on state, CI represses the production of LacI and TetR, keeping the system in that state, regardless of the subsequent levels of inducers.

repressing the first promoter (Fig. 3c). This could be realized by inserting the *cI* gene (as in the toggle switch) alongside *gfp* as a polycistron, and having the first promoter be repressed by CI (for example, the P_L promoter used in the toggle switch). In this case, once the system is switched to the on state by the simultaneous presence of IPTG and aTc, it will maintain this state regardless of the subsequent concentrations of inducers applied, because the expressed CI will repress the production of LacI and TetR. In this way, the system has memory such that the presence of the on state indicates that, at some point in the past, both IPTG and aTc were present simultaneously.

Recently, a new approach involving 'combinatorial synthesis' was used to generate a myriad of logical gene circuits[39]. This approach involved the clever use of subcloning and ligation, whereby 15 distinct promoter-gene units were constructed such that subsequent ligation of a mixture of the units yielded a library of three-gene networks. Specifically, the initial promoter-gene constructs incorporated uniquely designed *Bgl*I restriction sites in the polymerase chain reaction primers. This constrained the networks to the structure P_i-*lacI*-P_j-λ*cI*-P_k-*tetR*, where P_i, P_j and P_k were each one of the five promoters P^{L}_{1} (repressed by LacI), P^{L}_{2} (repressed by LacI), P^{T} (repressed by TetR), P^{λ} (repressed by CI) or P^{λ}_{2} (activated by CI). For measurement, a fourth transcriptional unit consisting of P^{λ}-*gfp* was incorporated in each plasmid, so that the input–output characteristics consisted of IPTG and aTc as inputs, and GFP fluorescence as output. The plasmid library was then transformed into *Escherichia coli* and grown under the four input conditions, with or without IPTG and with or without aTc. Analysis entailed the search for specific gene circuits in which the output fluorescence was a function of both inducers, and the result was a collection of logical circuits that included NAND, NOR and NOT IF gates.

Repressilator

Oscillations are used in engineering control systems as central 'clocks' to synchronize behaviour, and many multicellular organisms use a form of cellular 'clock' to coordinate their behaviour over the course of the day–night cycle[40,41]. These circadian rhythms manifest themselves in the periodic variation of concentrations of particular proteins in the cell. Although the precise molecular mechanism underlying even the most basic circadian rhythm is not fully understood, a number of general models describing these important rhythms have been proposed (refs 42–47, and see review in this issue by Goldbeter, pages 238–245).

When designing synthetic networks, an alternative to building a system that reproduces exact natural mechanisms is to generate systems that exhibit similar behaviour. This approach was used[14] to address the question of cellular oscillations, whereby a synthetic network (the 'repressilator') was produced that generated self-sustaining periodic oscillations in the concentrations of three proteins in a bacterial cell. The design operates on the same general principle as a ring oscillator in microelectronics. Accordingly, the repressilator

network architecture is cyclic (Fig. 4a), in which the LacI protein represses the promoter for the *tet* gene, the TetR protein represses the promoter for the *cI* gene, and the CI protein represses the promoter for the *lac* gene. As depicted in Fig. 4b, the network produced roughly sinusoidal oscillations in protein concentrations, observed by parallel expression of the reporter protein GFP.

As in the case of the toggle switch, a mathematical model was instrumental in the process of designing the repressilator. Although the ring network architecture is theoretically capable of sustaining oscillations[14,48], not all parameter choices give rise to oscillatory solutions. The modelling work indicated that oscillations were favoured by high protein synthesis and degradation rates, large cooperative binding effects, and efficient repression. These theoretical conclusions led to specific design choices: strong and tightly repressible hybrid promoters were selected, and the effective protein degradation rates were increased by *ssrA* tagging, whereby proteins are modified by the addition of an amino acid sequence which makes them targets for proteases in the cell.

An engineered circuit approach to sources of noise

Because the biochemical rates of transcription and translation are proportional to the number of promoter sites and messenger RNA molecules, these rates are typically small and imply relatively infrequent transcriptional and translational events compared with other interactions within the cell (for example, protein–protein interactions). In biochemistry, such infrequent events lead naturally to large fluctuations, and these fluctuations are known as internal noise because they originate from the underlying biochemical reactions rather than from some external perturbation or detection limitation.

The notion that such internal noise could be important in the choice of a developmental pathway for an organism has induced a flurry of modelling research devoted to the role of fluctuations in gene regulation (refs 25, 49–51; and see review in this issue by Arkin and co-workers, pages 231–237). Recently, theoretical models have been combined with engineered gene networks to elucidate the dominant source of internal noise in a single-gene network[20,23]. Given the two-step process of transcription and translation, the specific goal of this work was to determine their relative contribution to the fluctuations observed in the expressed protein concentrations within a cell.

Modelling work predicted that the random variation in expression from a single gene should scale linearly with the translational rate and be independent of the transcriptional rate[20]. Experimentally, point mutations were used to independently vary the transcriptional and translational rates, and the results were consistent with the theoretical predictions: the fluctuations in the expressed protein concentrations were observed to increase linearly with the translational efficiency while showing only a mild increase with the transcriptional efficiency[23]. Of particular note was the finding that the size of the fluctuations induced in the translational step was inversely proportional to the mRNA half-life, implying that fast mRNA turnover could be a means of mitigating noise. Because fast mRNA turnover increases the cellular energy requirement for protein production, the authors speculated that the evolution of gene regulation might entail a compromise between noise reduction and energy conservation.

The importance of tightly controlled amounts of cellular protein has led other researchers to model how specific network properties might act to decrease or utilize fluctuations[24,46,47,50,51]. One such study focused on a linear array of genes forming a network where each gene activates its nearest downstream neighbour[24]. The central finding was that cascades can act as attenuators for a noisy input signal, thus elucidating their potential importance in cell-wide signal transduction. Other studies showed how a model circadian network can function reliably in the presence of internal noise[46,47]. Although the underlying genetic architecture for the various known circadian systems has not been deduced, these networks seem to involve both positive- and negative-control elements[40]. This information was used to construct a generic model capable of oscillations that are resistant to fluctuations. The study also provided evidence that circadian oscillations might actually be enhanced by noise. This leads to the conjecture that the circadian circuitry has evolved to both reduce internal fluctuations and to exploit the residual noise that cannot be fully eliminated.

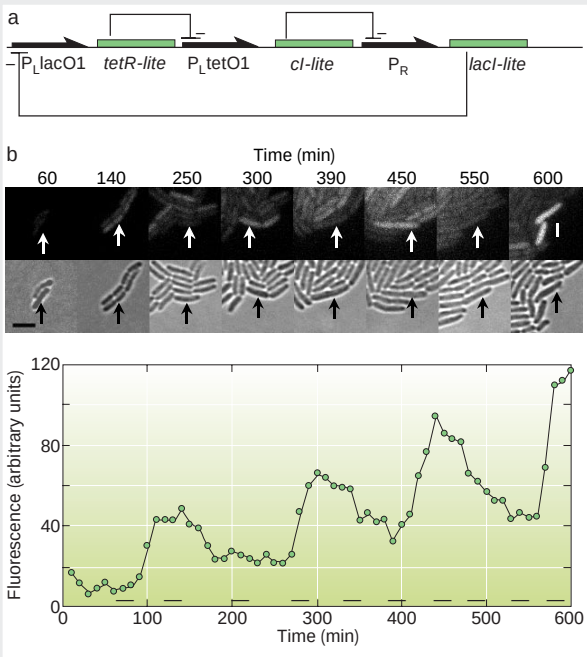

Figure 4 Synthetic transcriptional oscillator (the repressilator[14]). **a**, Network architecture. The synthetic system consists of three gene-promoter pairs arranged in a ring, such that each promoter's gene product represses the next promoter in the cycle. The promoter $P_L$lacO1 controls transcription of the gene *tetR-lite*, and the tetracycline repressor protein TetR represses the next promoter in the sequence, $P_L$tetO1. $P_L$tetO1 controls the transcription of *cI-lite*, and the protein CI represses the promoter P_R. Finally, P_R controls the expression of *lacI-lite*, and the lactose repressor protein LacI represses $P_L$lacO1, completing the cycle. Note that the suffix '*lite*' in the gene names refers to the presence of *ssrA* tags, which increase the degradation rate of the proteins. **b**, Experimental results showing oscillations in the repressilator[14]. The growth and timecourse of green fluorescent protein (GFP) expression was recorded for an individual *Escherichia coli* cell containing the repressilator plasmids; the cell was tracked using fluorescence (upper images) and bright-field (lower images) microscopy. Scale bar, 4 μm. The plot below these images shows a time series of fluorescence intensities, clearly indicating oscillatory behaviour in the cell. Bars at the bottom of the plot indicate the timing of cell division events. Note that the period of the oscillations is longer than the cell division time. (Adapted from ref. 14.)

The cascade and circadian network models described above provide clear theoretical predictions that can be tested systematically with engineered gene circuits. For example, cascades can be synthetically designed and the noise properties elucidated with single-cell microscopy. Similarly, the network underlying the proposed circadian oscillator could be built using an autocatalytic feedback loop as the primary network element[18].

Intercell signalling system

The use of signals to coordinate the behaviour of many individual devices is crucial in microelectronics and robotics, and cells also display a significant ability to communicate, both within multicellular organisms and within populations of unicellular organisms. Because cellular membranes act to isolate the cell from its environment, such communication generally relies on specialized chemicals that either pass through the membrane (through passive diffusion or active transport) or activate membrane-spanning receptors on the exterior of the cell.

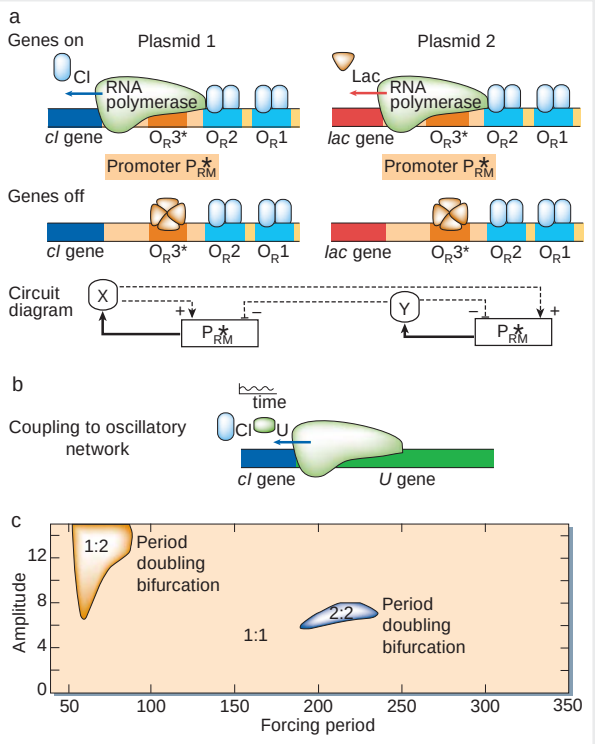

Figure 5 Synthetic oscillator design and synchronization properties. **a**, Schematic for the synthetic gene oscillator. The P^{*}_{RM} promoter is a mutant of the P_{RM} promoter that naturally exists in the virus λ phage[65]. In its natural state, the state of the virus is regulated by CI dimers, which bind to the three right operator sites O_R1, O_R2 and O_R3. In our design, the O_R3 operator is replaced with an operator region O_R3*, which has an affinity only for LacI (lactose repressor) tetramers. The depicted position of the Lac operator site is for illustrative purposes only, as the ideal placement of the operator may be upstream of O_R1 and O_R2. **b**, The synthetic oscillator is coupled to the host genome by inserting the *cI* gene adjacent to an oscillating gene product in the host. **c**, The resonance regions are depicted in a plot of the drive amplitude versus the drive frequency. Within these regions, the period of the synthetic network oscillations is entrained to that of the external drive. (Panels **a**–**c** redrawn from ref. 60.)

Experiments[52] have demonstrated the feasibility of sending signals between synthetic regulatory networks residing in different cells. This work made use of a well-studied, natural intercell signalling system, the quorum sensing pathway in the bacterium *Vibrio fischeri*[53,54]. In quorum sensing, bacteria regulate their behaviour based on the density of bacteria present nearby. Each bacterium secretes a signalling molecule (a homoserine lactone referred to as an 'autoinducer'), which passes through the cellular membrane in both directions. When many bacteria are present, the concentration of autoinducer reaches levels sufficient to activate a regulatory protein, LuxR, which then binds to the *lux* operator region and activates the expression of a suite of genes causing the bacterium to become luminescent.

In the synthetic system[52], two populations of cells were engineered: 'sender' cells containing an autoinducer synthase (LuxI) under the control of a chemically inducible promoter; and 'receiver' cells containing a reporter protein (GFP) controlled by the *lux* operator region. When the sender cells were induced to express LuxI, autoinducer was produced and diffused into the extracellular environment; the autoinducer then entered the receiver cells, and stimulated production of GFP by activating the *lux* region.

A simple example of coordinated behaviour is the synchronization of oscillators, and a recent modelling study[55] considered the use of the above-described intercell signalling system to synchronize a population of synthetic genetic relaxation oscillators[18]. Relaxation oscillators exhibit rapid transitions followed by periods of slow change, and previous theoretical work[56,57] showed that such oscillators are more readily synchronized than their more smoothly varying sinusoidal counterparts, such as the repressilator. The theoretical analysis indicated that rapid synchronization could be achieved by coupling each cell's production of autoinducer to its oscillatory phase. Experimentally verifying this prediction would be an interesting application of synthetic gene networks, as would a direct comparison of the synchronization behaviour of relaxation and sinusoidal genetic oscillators.

Applications

The above examples of engineered gene circuits serve to highlight how an integrated approach that combines computational modelling with experimental molecular biology can lead to insights into some of the basic modules that comprise complex, naturally occurring gene networks. The long-term goal of such work is to assemble increasingly complete models of the behaviour of natural systems, while maintaining at each stage the ability to test models in a tractable experimental system. An important complementary aspect of this approach is that the designer gene circuits which form the submodules will probably have important biotechnological applications in their own right. In this context, engineered gene networks represent a first step towards logical cellular control, whereby biological processes can be manipulated or monitored at the genetic level.

From the construction of a simple set of genetic building-block circuits (such as toggle switches and oscillators), one can imagine the design and construction of integrated biological circuits capable of performing increasingly elaborate functions. An integrated biological circuit could, like electronic control circuits, possess data-processing and storage circuitry, as well as input–output components necessary for sensing and affecting its environment. Ultimately, synthetic gene circuits encoded into DNA might be 'downloaded' into cells creating, in effect, a 'wet' nano-robot. These cellular robots could be used for a variety of functions, including *in vivo* biosensing, autonomously synthesizing complex biomaterials, executing programmed cell death, and interfacing with microelectronic circuits by transducing biochemical events to and from the electronics.

As an example of an integrated biological circuit, consider a recently engineered oncolytic adenovirus capable of selectively killing tumour cells[58]. In most tumour cells, the p53 gene network does not function properly, and this dysfunction leads to an unusually low amount of the tumour-suppressing p53 protein[59]. The engineered adenovirus is capable of detecting the presence or absence of p53, and executing a specific task depending on the p53 'state' of the cell. If the amount of p53 is normal, a viral promoter controlling the inhibition of replication is turned on and viral replication is halted. But if the amount of p53 is low, the virus detects the abnormal cell and replication proceeds along with the expression of viral proteins that lead to cell lysis and the spread of the adenovirus to other potentially cancerous cells.

Examples of other complex network-control schemes are provided by several recent modelling studies that focus on the utility of coupling designer gene networks to native cellular processes. One such study explored the coupling of an oscillating synthetic network to intrinsic cell-generated oscillations[60] (Fig. 5). This work provided design strategies for entraining and amplifying oscillations in cellular protein concentrations. Such control could prove useful in the design of networks that interact with cellular processes that require precise timing. Along these lines, seminal developments in the modelling of the cell-division cycle[61–63] could be coupled to the oscillator model, allowing for the design of protein delivery schemes that are signalled by the cellular growth cycle.

Another modelling study explored the utilization of engineered gene networks in the reverse engineering of large-scale gene

regulatory networks[64]. Here, the central idea is that small, engineered networks can be inserted into cells to provide a controlled perturbation mechanism for ongoing gene-expression experiments. One then tracks how the perturbation affects the genes in a naturally occurring network, and this information can be used to deduce the network topology. This method may prove useful in identifying and validating specific drug targets and in unravelling the effects of chemical compounds.

By reducing the complexity of the systems under study, synthetic gene networks offer the ability to gain a detailed understanding of the mechanisms involved in gene regulation. As our grasp of the fundamental principles of gene regulation improves, we will be able to design and study increasingly complex systems. The conjunction of the advanced experimental techniques of molecular biology with the mathematical tools of nonlinear dynamics and statistical physics provides an exciting opportunity for rapid advances in the understanding and control of cellular behaviour. □

doi:10.1038/nature01257

1. Monod, J. & Jacob, F. General conclusions: telenomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26,** 389–401 (1961).
2. Glass, L. & Kauffman, S. A. The logical analysis of continuous, non-linear biochemical control networks. *J. Theor. Biol.* **39,** 103–129 (1973).
3. Savageau, M. A. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252,** 546–549 (1974).
4. Kauffman, S. A. The large-scale structure and dynamics of gene control circuits: an ensemble approach. *J. Theor. Biol.* **44,** 167–190 (1974).
5. Glass, L. Classification of biological networks by their qualitative dynamics. *J. Theor. Biol.* **54,** 85–107 (1975).
6. Glass, L. Combinatorical and topological methods in nonlinear chemical kinetics. *J. Chem. Phys.* **63,** 1325–1335 (1975).
7. Savageau, M. A. *Biochemical System Analysis* (Addison Wesley, Reading, 1976).
8. Goodwin, B. C. *Analytical Physiology of Cells and Developing Organisms* (Academic, London, 1976).
9. Tyson, J. J. & Othmer, H. G. The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* **5,** 1–62 (1978).
10. McAdams, H. H. & Shapiro, L. Circuit simulation of genetic networks. *Science* **269,** 650–656 (1995).
11. McAdams, H. H. & Arkin, A. Towards a circuit engineering discipline. *Curr. Biol.* **10,** R318–R320 (2000).
12. Reinitz, J. & Vaisnys, J. R. Theoretical and experimental analysis of the phage lambda genetic switch implies missing levels of co-operativity. *J. Theor. Biol.* **145,** 295–318 (1990).
13. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403,** 339–342 (2000).
14. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).
15. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405,** 590–593 (2000).
16. Hasty, J., Pradines, J., Dolnik, M. & Collins, J. J. Noise-based switches and amplifiers for gene expression. *Proc. Natl Acad. Sci. USA* **97,** 2075–2080 (2000).
17. Smolen, P., Baxter, D. A. & Byrne, J. H. Mathematical modeling of gene networks. *Neuron* **26,** 567–580 (2000).
18. Hasty, J., Isaacs, F., Dolnik, M., McMillen, D. & Collins, J. J. Designer gene networks: towards fundamental cellular control. *Chaos* **11,** 207–220 (2001).
19. Hasty, J., McMillen, D., Isaacs, F. & Collins, J. J. Computational studies of gene regulatory networks: *in numero* molecular biology. *Nature Rev. Genet.* **2,** 268–279 (2001).
20. Thattai, M. & van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl Acad. Sci. USA* **98,** 8614–8619 (2001).
21. Simpson, M. L., Sayler, G. S., Fleming, J. T. & Applegate, B. Whole-cell biocomputing. *Trends Biotechnol.* **19,** 317–323 (2001).
22. Becskei, A., Séraphin. B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20,** 2528–2535 (2001).
23. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
24. Thattai, M. & van Oudenaarden, A. Attenuation of noise in ultrasensitive signaling cascades. *Biophys. J.* **82,** 2943–2950 (2002).
25. McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA* **94,** 814–819 (1997).
26. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402,** C47–C51 (1999).
27. Lauffenburger, D. A. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl Acad. Sci. USA* **97,** 5031–5033 (2000).
28. Palsson, B. O. & Lightfoot, E. N. Mathematical modeling of dynamics and control in metabolic networks. *J. Theor. Biol.* **113,** 279–298 (1985).
29. Thomas, R. The role of feedback circuits: positive feedback circuits are a necessary condition for positive real eigenvalues of the Jacobian matrix, *Ber. Besenges. Phys. Chem.* **98,** 1148–1151 (1994).
30. Thomas, R., Thieffry, D. & Kaufman, M., Dynamical behaviour of biological regulatory networks-I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull. Math. Biol.* **57,** 247–276 (1995).
31. Ferrell, J. E. Jr Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr. Opin. Cell. Biol.* **14,** 140–148 (2002).
32. Keller, A. Model genetic circuits encoding autoregulatory transcription factors. *J. Theor. Biol.* **172,** 169–185 (1995).
33. Ferrell, J. E. Jr Xenopus oocyte maturation: new lessons from a good egg. *BioEssays* **21,** 833–842 (1999).
34. Ferrell, J. E. Jr Building a cellular switch: more lessons from a good egg. *BioEssays* **21,** 866–870 (1999).
35. Aurell, E., Brown, S., Johanson, J. & Sneppen, K. Stability puzzles in phage λ. *Phys. Rev. E* **65,** 051914-1–051914-9 (2002).
36. Bialek, W. in *Advances in Neural Information Processing* Vol. 13 (eds Leen, T. K., Dietterich, T. G. & Tresp, V.) 103–109 (MIT Press, Cambridge, MA, 2001).
37. Weiss, R. & Basu, S. The device physics of cellular logic gates. First Workshop on Non-Silicon Computing <http://www-2.cs.cmu.edu/~phoenix/nsc1/paper/3-2.pdf> (2002).
38. Weiss, R. *Cellular Computation and Communications Using Engineered Genetic Regulatory Networks.* Thesis, Massachusetts Institute of Technology (2001).
39. Guet, C., Elowitz, M., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296,** 1466–1470 (2002).
40. Dunlap, J. Molecular bases for circadian clocks. *Cell* **96,** 271–290 (1999).
41. Panda, S., Hogenesch, J. B. & Kay, S. A. Circadian rhythms from flies to human. *Nature* **417,** 329–335 (2002).
42. Leloup, J. C. & Goldbeter, A. A model for circadian rhythms in *Drosophila* incorporating the formation of a complex between the PER and TIM proteins. *J. Biol. Rhythms* **13,** 70–87 (1998).
43. Tyson, J. J., Hong, C. I., Thron, C. D. & Novak, B. A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. *Biophys. J.* **77,** 2411–2417 (1999).
44. Leloup, J. C. & Goldbeter, A., Modeling the molecular regulatory mechanism of circadian rhythms in *Drosophila*. *BioEssays* **22,** 84–93 (2000).
45. Roussel, M. R., Gonze, D. & Goldbeter, A. Modeling the differential fitness of cyanobacterial strains whose circadian oscillators have different free-running periods: comparing the mutual inhibition and substrate depletion hypotheses. *J. Theor. Biol.* **205,** 321–340 (2000).
46. Barkai, N. & Leibler, S. Biological rhythms: circadian clocks limited by noise. *Nature* **403,** 267–268 (2000).
47. Vilar, J. M. G., Kueh, H. Y., Barkai, N. & Leibler, S. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl Acad. Sci. USA* **99,** 5988–5992 (2002).
48. Smith, H. Oscillations and multiple steady states in a cyclic gene model with repression. *J. Math. Biol.* **25,** 169–190 (1987).
49. Arkin, A., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* **149,** 1633–1648 (1998).
50. Kepler, T. & Elston, T. Stochasticity in transcriptional regulation: origins, consequences and mathematical representations. *Biophys. J.* **81,** 3116–3136 (2001).
51. Paulsson, J., Berg, O. G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl Acad. Sci. USA* **97,** 7148–7153 (2000).
52. Weiss, R. & Knight, T. F. in *DNA6: Sixth International Meeting on DNA Based Computers* (Leiden, The Netherlands, 2000).
53. Fuqua, C., Winans, S. & Greenberg, E. P. Census and consensus in bacterial ecosystems: the LuxR-LuxI family of transcriptional regulators. *Annu. Rev. Microbiol.* **50,** 727–751 (1996).
54. Bassler, B. L. How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr. Opin. Microbiol.* **2,** 582–587 (1999).
55. McMillen, D., Kopell, N., Hasty, J. & Collins, J. J., Synchronizing genetic relaxation oscillators by intercell signaling. *Proc. Natl Acad. Sci. USA* **99,** 679–684 (2002).
56. Somers, D. & Kopell, N. Rapid synchronization through fast threshold modulation. *Biol. Cybern.* **68,** 393–407 (1993).
57. Wang, D. L. in *Proc. 15th Annu. Conf. Cognit. Sci. Soc.* 1058–1063 (Lawrence Erlbaum Assoc., Hillsdale, NJ, 1993).
58. Ramachandra, M. *et al.* Re-engineering adenovirus regulatory pathways to enhance oncolytic specificity and efficacy. *Nature Biotech.* **19,** 1035–1041 (2001).
59. Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* **408,** 307–310 (2000).
60. Hasty, J., Dolnik, M., Rottschafer, V. & Collins, J. J. A synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* **88,** 148101-1–148101-4 (2002).
61. Novak, B. & Tyson, J. J. Modeling the control of DNA replication in fission yeast. *Proc. Natl Acad. Sci. USA* **94,** 9147–9152 (1997).
62. Sveiczer, A., Csikasz-Nagy, A., Gyorffy, B., Tyson, J. J. & Novak, B. Modeling the fission yeast cell cycle: quantized cycle times in *wee1⁻cdc25Δ* mutant cells. *Proc. Natl Acad. Sci. USA* **97,** 7865–7870 (2000).
63. Chen, K. C., Csikasz-Nagy, A., Gyorffy, B., Novak, B. & Tyson, J. J. Kinetic analysis of a molecular model of the budding yeast cell cycle. *Mol. Biol. Cell* **11,** 369–391 (2000).
64. Tegner, J., Yeung, M. K. S., Hasty, J. & Collins, J. J. Reverse engineering gene networks—integrating genetic perturbations with dynamical modeling. *Proc. Natl Acad. Sci. USA* (submitted).
65. Ptashne, M. *et al.* How the λ repressor and cro work. *Cell* **19,** 1–11 (1980).

Acknowledgements

This work was supported by the Defense Advanced Research Projects Agency (DARPA), Office of Naval Research (ONR), National Science Foundation (NSF) BioQuBIC, the Fetzer Institute, and Natural Sciences and Engineering Research Council of Canada (NSERC).