Annu. Rev. Biomed. Eng. 2003. 5:179–206
doi: 10.1146/annurev.bioeng.5.040202.121553

THE ENGINEERING OF GENE REGULATORY NETWORKS

Mads Kærn, William J. Blake, and J.J. Collins
Center for BioDynamics, Department of Biomedical Engineering, and Bioinformatics Program, Boston University, Boston, Massachusetts 02215; email: mkaern@bu.edu

Key Words circuits, switches, modules, noise, models

■ **Abstract** The rapid accumulation of genetic information and advancement of experimental techniques have opened a new frontier in biomedical engineering. With the availability of well-characterized components from natural gene networks, the stage has been set for the engineering of artificial gene regulatory networks with sophisticated computational and functional capabilities. In these efforts, the ability to construct, analyze, and interpret qualitative and quantitative models is becoming increasingly important. In this review, we consider the current state of gene network engineering from a combined experimental and modeling perspective. We discuss how networks with increased complexity are being constructed from simple modular components and how quantitative deterministic and stochastic modeling of these modules may provide the foundation for accurate in silico representations of gene regulatory network function in vivo.

CONTENTS

1. INTRODUCTION

Cells are complex processors of information, able to integrate and respond to multiple signals in a robust manner. The mechanisms by which cells are able to achieve this in a constantly fluctuating intra- and extracellular environment are remarkable and many. Current research is heavily focused on dissecting the

regulatory circuitry of the cell, and some have proposed a modular approach to such study (1, 2). The high degree of complexity involved in cellular response can be simplified by considering large-scale genetic networks as composed of subsets of simpler components, or modules (3, 4), that are interconnected through input and output signals, analogous to electrical circuits (5, 6). Interestingly, the parallels between electrical and genetic circuit engineering extend beyond the descriptive level. In the same way that electrical engineers construct circuits, genetic network engineers make use of the biological equivalents of inverters and transistors to manipulate living organisms by connecting these modules into gene regulatory networks that can control cellular function.

In this review, we discuss how artificial gene networks can be constructed from well-characterized natural components. Over the past few years, simple genetic inverters and transistors have been combined into circuits that have memory (7–9), logic gate functionality (10, 11), and oscillatory dynamics (12). In the years to come, it can be anticipated that artificial gene networks will become more sophisticated, allowing the precise and multifaceted control of cellular function. As the number of interconnected modules in a network increases, the tools to interpret and analyze their function will become increasingly important [see, e.g., (13)]. Here, we discuss several methods currently being used in qualitative and quantitative modeling of gene networks. Rather than discussing how mathematics is used to describe gene regulatory systems in general [for recent reviews, see (14–18)], we focus on the development of accurate mathematical representations of individual genetic modules and review the modeling tools required for this purpose. We propose a combined experimental and theoretical approach to the construction of novel gene networks where well-characterized modules are quantitatively described and rationally assembled into more complex circuits based on the predictions of quantitative models.

2. NATURAL MODULES: COMPONENT PARTS

Natural gene networks can be described as circuits of interconnected functional modules, each consisting of specialized interactions between proteins, DNA, RNA, and small molecules. In this review, we define a module as the simplest element of a gene regulatory network, consisting of a promoter, the gene(s) expressed from that promoter, and the regulatory proteins (and their cognate DNA binding sites) that affect the expression of that gene. We address these modules as logic elements (19), where the output is the level of gene expression (measured by the amount of mRNA or protein produced) and the inputs are factors (e.g., regulatory proteins and small molecules) that can affect the expression of the given gene.

Generally, a promoter describes the region of DNA to which the RNA polymerase binds prior to the transcription of downstream gene(s). Specific sequences of DNA within or near promoter elements serve as binding sites for transfactors, which can either increase (activators) or decrease (repressors) the probability that the gene is transcribed into mRNA. Transactivators commonly act to enhance the

initiation of transcription through, for example, recruitment of RNA polymerase to the promoter, whereas many transrepressors act to interfere with transcriptional initiation [for a recent review, see (20)]. Below, we discuss three natural modules and their component parts that are widely used in the engineering of artificial gene regulatory networks.

In prokaryotic cells, multiple genes involved in a specific biological process (e.g., metabolism of a certain carbon source) are coordinately expressed from a single promoter in a natural genetic logic module called an operon. The *lac* operon in *Escherichia coli*, which is one of the best characterized systems of prokaryotic gene regulation (21–23), is composed of three genes (*lacZ*, *lacY*, and *lacA*) involved in the metabolism and uptake of the disaccharide lactose. The output level of these genes is controlled by binding of the Lac repressor (LacI) to the *lac* operator site (*lacO*) within the *lac* operon promoter, P_{lac}. In the absence of lactose, expression from P_{lac} is inhibited by the binding of LacI to *lacO*, which efficiently represses expression of the *lac* operon by interfering with the binding of RNA polymerase to the promoter region [for details, see, e.g., (22–26)]. Lactose (more specifically, allolactose), or the chemical analog isopropylthio-β-galactoside (IPTG), decreases the affinity of the repressor for its operator site and induces expression of the *lacZYA* genes. Notably, one of the induced genes, *lacY*, codes for a constituent of the membrane-bound system transporting lactose into the cell, making the *lac* operon a classic example of a genetic switch relying on autocatalytic feedback (27). In addition, the P_{lac} promoter can be viewed as the biological equivalent of an inverter because increased activity of LacI (the input) causes a decrease in expression from P_{lac} (the output). It may also be viewed as a transistor because the expression is effectively modulated by inducer (the input) when *lacI* is constitutively expressed. Components from well-characterized natural gene networks, such as the *lac* operon, are ideal candidates for artificial gene network construction.

A second bacterial repressor-operator system in broad use for constructing artificial gene networks is the Tet system derived from the transposon Tn10 [for a comprehensive review of this system, see (28)]. The Tet system confers bacterial resistance to tetracycline antibiotics, which interfere with protein synthesis on ribosomes. Two genes, *tetR* and *tetA*, modulate tetracycline resistance in a manner similar to the functioning of the *lac* operon. Specifically, in the absence of tetracycline (Tc), the Tet repressor (TetR) binds to *tet* operator sites within the promoter controlling the expression of *tetA*, an antiporter, which effectively drives the efflux of Tc from the cell (29). Binding of Tc to TetR decreases the affinity of the repressor to *tetO*, causing upregulation of *tetA* expression and subsequent removal of Tc from the cell. The interactions between TetR, *tetO*, and inducer have been extensively studied [e.g., see (30–34)], and these components are used ubiquitously in the engineering of gene regulatory networks.

A third natural module involves a regulatory protein, the λ repressor, which acts to both negatively and positively regulate gene expression. The λ repressor is part of a natural genetic switch that controls the lysis/lysogeny decision of bacteriophage

λ after infection of a bacterial cell (35). This process has been studied in detail at the molecular level (36), and here we describe only the components used in the engineering of artificial gene networks. The λ repressor is produced from the *cI* gene expressed from the P_{RM} promoter. Its production is autoregulated due to feedback binding of the λ repressor to three operators, denoted OR1, OR2, and OR3. Binding of the λ repressor to OR1 does not directly affect expression from P_{RM} but increases the binding affinity of λ repressor to OR2. Subsequent binding of λ repressor to OR2 activates expression from P_{RM}. When the concentration of λ repressor is increased further, it binds to OR3, which causes the downregulation of expression from P_{RM}. The λ repressor also modulates expression from the P_R promoter, which partially overlaps with P_{RM}, and represses expression from a third promoter denoted P_L.

3. MODULE ENGINEERING: ARTIFICIAL GENE NETWORKS

3.1. Simple Networks

We have introduced modularity in terms of simple logic units involving a single promoter and the downstream gene(s), where regulatory protein and inducer inputs affect the level of protein output(s). In fact, regulatory proteins (37) and their binding sites are also modular in that different domains from different proteins can be combined to yield hybrid proteins of novel function [e.g., see (38)]. Similarly, the binding sites of activators and repressors can be inserted into promoter elements to allow for novel transcriptional control of gene expression from hybrid promoters.

An inducible switch, as illustrated in Figure 1*a*, is a simple engineered gene network that essentially mimics the natural bacterial operon. In general, this switch

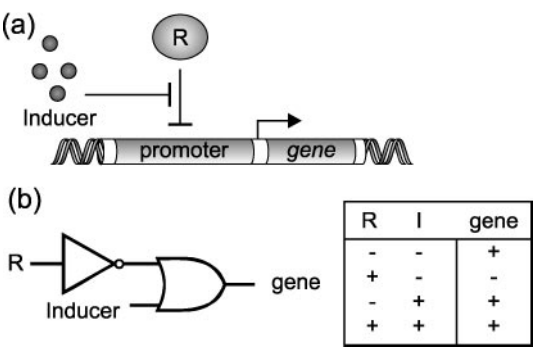

Figure 1 Architecture of the inducible switch gene network. (*a*) The promoter contains operator(s) for a repressor (R) whose activity is modulated by a chemical inducer (I). The level of repressor protein is controlled by a second constitutively expressing promoter (not shown). (*b*) The circuit diagram and boolean logic of the inducible switch.

is constructed by "addressing" a repressor protein to an otherwise constitutive promoter by insertion of the repressor's DNA binding site (or operator) at an appropriate position within or near the promoter sequence. In the presence of the repressor protein (R), the switch is in the OFF position, and expression of the downstream gene is inhibited. By adding inducer (I), the switch is moved to the ON position and expression of the downstream gene is activated. Expression from the promoter is thus amplified when the amount of inducer increases. As described above for the P_{lac} module, the inducible switch acts as a logic gate and can be represented by the circuit diagram and boolean logic shown in Figure 1*b*.

Two commonly used inducible switches were constructed in *E. coli* (39) by inserting tandem operators for TetR or LacI into a modified P_{L} promoter to make the hybrid promoters $P_{\mathrm{LtetO-1}}$ and $P_{\mathrm{LlacO-1}}$, respectively. Expression from these hybrid promoters is tightly repressed by TetR or LacI and can be modulated over a broad dynamic range by tuning the amounts of the inducer Tc (or the nontoxic chemical analog, anhydrotetracycline, ATc) or IPTG. A third promoter, designated $P_{\mathrm{lac/ara-1}}$, developed by the same group (39), contains the binding site for the activating transfactor synthesized from the *araC* gene, as well as the *lac* operator. This module has four inputs, the proteins LacI and AraC and their respective inducers IPTG and arabinose, and is a demonstration of how individual modules can be engineered to accept multiple inputs.

The importance of the modular components described above is exemplified by their use in networks that can function in cells of higher complexity. Specifically, prokaryotic components, such as repressors and their operator sites, can be transplanted into eukaryotic cells without loss of function or binding specificity (38, 40, 41). An additional advantage of using prokaryotic elements in eukaryotic cells is the avoidance of unwanted interference with the expression of nontargeted genes (pleiotropic effects) (41, 42). Inducible switch systems based on the Tet or Lac repressor have been developed in a large number of organisms [e.g., see (43–58)], setting the stage for the construction of more complex gene networks in a broad range of biological systems.

In contrast to most prokaryotic genes, eukaryotic genes are generally in a silenced state (59) due to the organization of genetic material in tightly packaged structures called chromatin. Therefore, it is common for many eukaryotic genes to require some mechanism of activation before they can be expressed (60). An elegant demonstration of the versatility of genetic module engineering is the construction of novel transfactor proteins that can activate gene expression in mammalian cells (53, 54). In these studies, novel transactivators were constructed by fusing the DNA binding domain of the Tet repressor to the transactivation domain of the protein VP16 from the *Herpes simplex* virus. Insertion of seven copies of the *tet* operator ($7 \times tetO$) upstream of the P_{CMV} promoter region allowed tight control of downstream gene expression. Two variants that have differential responses to inducer were developed. In the TetOFF system, illustrated in Figure 2*a*, the tetracycline-dependent transactivator (tTA) binds to the *tet* operators in the absence of Tc, activating transcription from the promoter. Conversely, in the TetON

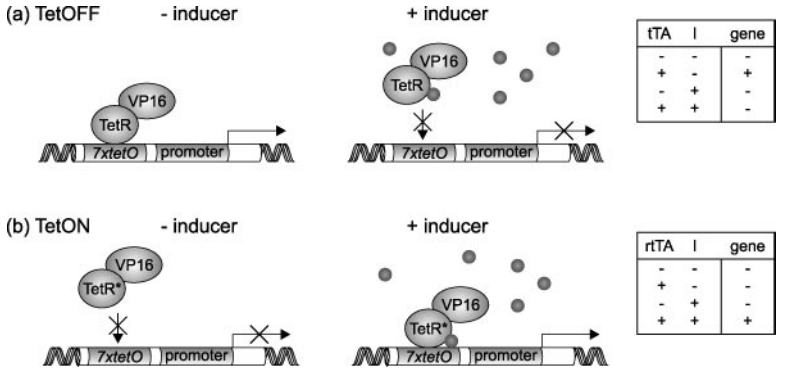

Figure 2 Eukaryotic tetracycline-dependent transactivator systems involving fusion transactivators comprised of the DNA binding domain of the TetR repressor and the activation domain of the viral VP16 transactivator protein. (*a*) The TetOFF system. The fusion protein TetR-VP16 (tTA) binds to 7 × *tetO* and activates transcription. The inducer (Tc or Tc analogues) attenuates expression by preventing binding of tTA to *tetO*. (*b*) The TetON system. The fusion protein variant TetR*-VP16 (rtTA) binds to *tetO* and activates expression only in the presence of inducer.

system, illustrated in Figure 2*b*, the reverse tTA (rtTA) binds and activates expression in the presence of Tc.

Successful construction of a variety of artificial networks requires that the number of characterized components is increased beyond those described above. To this end, researchers have developed alternative systems based on other repressor-operator combinations. For example, repressible (PipOFF) and inducible (PipON) systems were developed (61) based on the Pip (pristinamycin-induced protein) repressor of *Streptomyces coelicolor* for the control of gene expression in mammalian cells. Additionally, similar repressible (E_{OFF}) and inducible (E_{ON}) systems (62) were recently developed for controlling mammalian gene expression based on the repressor-operator system of the *E. coli* erythromycin-resistance regulon. Numerous research groups have constructed and implemented various other gene expression systems, including a light-switchable gene system (63), and the literature is far too extensive to be covered here. We refer the interested reader to a comprehensive review on artificial mammalian gene networks (64).

3.2. Complex Networks

The *tetR*, *lacI*, and *cI* modules described in Section 3.1 have been used to construct a variety of complex gene networks in *E. coli*. The bacterial toggle switch (7) and the bacterial ring oscillator, or repressilator (12), illustrated in Figures 3*a* and 3*b*, respectively, were constructed by connecting the inputs and outputs of the *tetR*, *lacI*, and *cI* repressible modules discussed above. As illustrated in Figure 3*a*, the toggle switch is composed of two co-repressive genes, *rep1* and *rep2*, producing

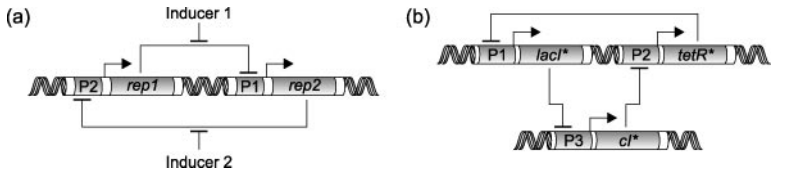

Figure 3 Complex networks engineered in *E. coli.* (*a*) The toggle switch. The protein encoded by the *rep1* gene represses the expression from promoter P1. The protein encoded by *rep2* represses the expression from promoter P1. This mutual repression enables the system to be in one of two different expression states and switch between these states by transient addition of inducer 1 or inducer 2. (*b*) The ring oscillator constructed from three serially connected TetR, LacI, and λ repressor-based switch modules. $P1 = P_{\mathrm{LtetO-1}}$, $P2 = P_{\mathrm{R}}$, $P3 = P_{\mathrm{LlacO-1}}$. The oscillator was constructed using rapidly degraded variants (denoted *) of the repressor proteins.

a system with two states (Rep1 ON/Rep2 OFF and Rep1 OFF/Rep2 ON). It was implemented in two versions using either LacI and TetR (designated pIKE) or LacI and the λ repressor (designated pTAK). The *cI* gene used in the pTAK system was a mutant version termed *cI857*, which produces a λ repressor protein that is inactivated at elevated temperatures (65). Transient pulses of IPTG or ATc (pIKE system), or of IPTG or high temperature (pTAK system), cause robust switching between states. The state of the pTAK toggle switch, as well as two other independently developed LacI/λ repressor-based switches (8, 9), has shown stable transfer over successive generations, and the toggle constitutes a system that allows nongenetic inheritable (epigenetic) information and retains memory of, and properties determined by, events in the ancestry of the cell.

The bacterial ring oscillator was constructed by connecting three repressible modules in series (12). The Tet repressor was expressed from the $P_{\mathrm{LlacO-1}}$ promoter, the Lac repressor from P_{R}, and the λ repressor from the $P_{\mathrm{LtetO-1}}$ promoter, thus constituting a closed ring with negative feedback to the previous module as illustrated in Figure 3*b*. The system showed a sinusoidal oscillation with a period of approximately 2.5 h. An oscillator with a tunable period and amplitude would be extremely useful to probe and characterize gene circuits in a manner similar to the way in which signal generators are used for system identification and transfer function characterization in electrical engineering. Unfortunately, the ring oscillator construct behaves somewhat erratically and cells oscillate without phase coherence with a period of 160 $\pm$ 40 min in only 40% of the cells (12). It can be speculated that these problems may be addressed by the construction of a more noise-robust relaxation oscillator (66, 67), which, when coupled to an engineered cell-to-cell communication system (68), can be synchronized across the population (69).

The toggle switch and ring oscillator demonstrate how appropriately adjusted individual network modules can be connected in a rational manner to produce a desired output. An alternative approach is to connect various modules at random,

thereby generating a library of all the possible network connectivities. Individual networks generated in this way can then be screened for the desired functionality and isolated for further optimization. This approach was recently taken to generate a subset of the 125 possible networks that can be constructed with the *lacI*, *tetR*, and *cI* genes and five different promoters responding in various ways to each repressor protein (10). In addition to the $P_{\mathrm{LlacO-1}}$, $P_{\mathrm{LtetO-1}}$, and P_{R} promoters used to construct the ring oscillator, the study employed the λ repressor-activated P_{RM} promoter and an additional LacI-repressed promoter. The networks were constructed by first generating the 15 possible promoter/gene combinations and then randomly connecting these DNA fragments in such a way that each network contained one copy of the *lacI*, *tetR*, or *cI* genes. The randomly generated networks were transformed into *E. coli* and their responses to ATc and IPTG were determined. This screen revealed the logic gate phenotypes NOR, NOT IF, and NAND. Cells with a variety of responses were isolated, and sequencing of their DNA content revealed 13 different network topologies that included autoregulatory circuits, regulatory cascades with and without feedback, as well as a number of toggle switches.

Connecting simple modules at random (10) is an elegant demonstration of the combinatorial diversity in gene networks and complements the more rational approach of connecting specific modules for a desired outcome. In general, it can be expected that the kinetic properties of individual modules must be appropriately adjusted to obtain a specific network function. Rather than using intuition and trial-and-error-based approaches once the network has been assembled, it may be beneficial to rationally manipulate individual modules before connecting them together. Although such standardized characterization of gene network modules is time-consuming, it will provide the necessary information for the systematic improvement of network modules and for the model-assisted design of complex gene networks.

The characterization of a network module requires careful measurement of the relationship between the module inputs and output. A genetic cascading circuit for the standardized characterization of network modules is illustrated in Figure 4*a*. It allows indirect measurement of inputs (primarily transfactors), which are difficult to measure directly. Outputs can be measured directly using a reporter gene, such as green fluorescent protein (GFP). The cascading expression system in Figure 4*a* was used in *E. coli* (11) to characterize a logic module comprising a repressor Rep2 (LacI or λ repressor) and a promoter, P2, containing binding sites for Rep2 (*lacO* or operators for the λ repressor). In this circuit, the transfactor input signal is controlled from an inducible switch comprising the constitutively expressed repressor, Rep1, and the promoter P1. The level of expression of *rep2* can therefore be set by the addition of inducer 1. The input signal (the amount of the Rep2 protein) is measured indirectly by coexpressing the cyan fluorescent reporter (CFP) from promoter P1, whereas the output of the Rep2/P2 module is measured by expressing the yellow fluorescent reporter (YFP) from the P2 promoter. The direct measurement of the input/output relationship allows immediate access to critical information regarding the function of the module, such as basal expression level, dynamic range,

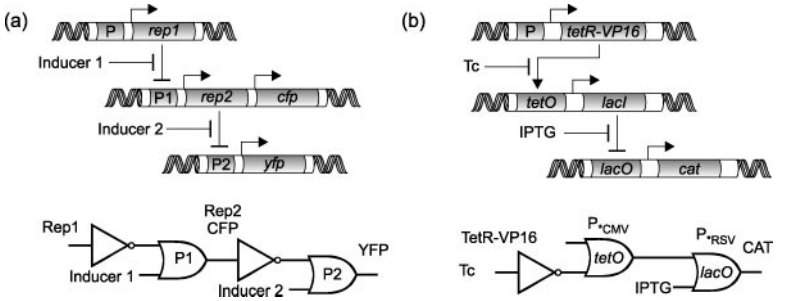

Figure 4 Genetic cascading circuits. (*a*) Prokaryotic system used to characterize the response of the Rep2/P2 repressor/promoter switch module to the inputs Rep2 (repressor protein) and inducer 2. The expression of *rep2* is regulated from the Rep1/P1 switch module by addition of inducer 1. The relation between Rep2 (the input) and the expression from P2 (the output) is measured by coexpression of CFP from P1 and by expressing YFP from P2. (*b*) Mammalian system allowing multifaceted control of gene expression using the rTA module and a mammalian inducible Lac switch module. Addition of inducer (Tc) allows TetR-VP16-activated expression of *lacI* from the *tetO*-containing promoter. LacI, in turn, represses expression of the *cat* gene from the *lacO*-containing promoter. The latter step can be modulated by addition of IPTG.

response coefficients, control strength, and effective Hill coefficients. Knowledge of such parameters is essential for the success of network design guided by quantitative description of individual network modules, as is discussed in Section 4.3. Once implemented, standardized module characterization will also allow a more rapid procedure for adjusting the kinetic properties of individual network modules and offer a simple way to systematically test how genetic manipulations affect the properties of a given module.

A cascading circuit somewhat similar to the bacterial version illustrated in Figure 4*a* has been constructed and implemented in mammalian cells (70). Although this system would be ideal for module characterization, it was constructed with the purpose of tightly controlling gene expression. As illustrated in Figure 4*b*, the network is based on two of the simpler modules described in Section 3.1: the tTA module and the Lac repressor module. Constitutive expression of tTA activates transcription from the *tetO*-containing $P_{*\mathrm{CMV}}$ promoter. LacI, expressed from $P_{*\mathrm{CMV}}$, binds to *lac* operator sites within the modified $P_{*\mathrm{RSV}}$ promoter, thereby downregulating the expression of the CAT (Cm acetyltransferease) reporter. The level of *cat* gene expression, or of any other gene inserted at this location, is thus subject to multifaceted control. Addition of Tc lowers the amount of LacI, and expression from the *lacO*-containing $P_{*\mathrm{RSV}}$ promoter can be further attenuated with IPTG. This cascading circuit is an important proof-of-principle demonstration that complex regulatory gene networks allowing multifaceted control can indeed be engineered into higher organisms.

4. MODELING ENGINEERED GENE NETWORKS

As the complexity of engineered networks increases, the ability to predict their behavior will become increasingly important. In this section, we discuss a number of complementary approaches that can be used to model gene networks. Rather than focusing on the general approaches, we consider only the mathematical descriptions of individual network modules that we think have the greatest potential for the engineering of complex gene circuits. The goal is to use such descriptions to accurately predict the properties and the function of modules connected into networks, and to make in silico suggestions for optimal design strategies prior to implementation in vivo.

In Section 4.1, we discuss phenomenological and mass action kinetic models that are based on ordinary differential equations (ODEs). When applied to simple circuits, models of this type are used to approximate the average level of gene expression within populations of genetically identical cells, or the average level of gene expression within single cells measured over a long time period. For more complex circuits, phenomenological and mass action kinetic models can be used to predict the dynamics of a network and to explore how modules should be connected or modified to achieve a desired network functionality. These types of models offer simplicity in analysis and interpretation at the expense of capturing behavior of individual cells and the resulting heterogeneity within a population of cells.

Population heterogeneity arises, in part, from stochasticity in gene expression, or gene expression noise. In Section 4.2, we consider simple models of gene expression that take into consideration the inherent stochastic nature of chemical reactions. The traditional mass action kinetics, discussed in Section 4.1, was developed for "test tube" chemical systems, and its applicability requires that a number of conditions are fulfilled [see, e.g., (71)]. The approach does not address the effects of "quantum biochemistry" (72) arising from reactions that involve molecules, such as promoter elements and operator sites, that are present in very few copies within a cell. The effects of the probabilistic nature of individual reaction events become more pronounced as the number of molecules decreases, and the deviation from deterministic behavior may be an important factor affecting the properties of individual cells. Of particular interest in the present engineering context is the use of relatively simple stochastic models to deduce how gene expression noise can be manipulated experimentally to improve network function by increasing the signal-to-noise ratio within cell populations.

In Section 4.3, we discuss an approach that combines stochastic, deterministic, phenomenological, and mass action–based modeling. When combined with data obtained from experimental module characterization and reaction mechanisms determined by more traditional means, this approach can provide an accurate description of how individual network modules function within individual cells. We think that such an approach can be used to construct accurate mathematical descriptions of individual network module function, which, in turn, will enable model-based, rational engineering of complex gene networks.

4.1. Deterministic Models

Most current models of gene networks are formulated in terms of deterministic ODEs and rely on traditional mass action kinetics or on phenomenological representations of reaction mechanisms. The construction of the bacterial toggle switch (7) is an excellent example of how a purely phenomenological deterministic model can be used to assist the construction of engineered gene networks. The model is based on simple descriptions of inducible switches, such as the Tet and Lac switches described in Section 3.1, and assumes that the dependence between the rate (ν) of expression from the switch module and the amounts of repressor (R) and inducer (I) is given by

$$\nu = \frac{\nu_{max}}{K_m + R_a^{\gamma}}, \quad R_a = \frac{R}{(1 + I/K_I)^{\mu}}, \tag{1}$$

where R_a is the number of active repressor molecules, K_m and K_I are equilibrium constants, and ν_{max} is the maximal rate of expression (obtained in the absence of repressor, or at full induction). The coefficients γ and μ determine how the rate of expression and repressor activity are attenuated in the presence of repressor and inducer, respectively.

The toggle model presented in Reference 7 is a combination of two inducible switch models and describes the evolution of the concentrations of two repressor proteins, denoted by u for LacI and v for the λ repressor (the pTAK toggle), in terms of simple ODEs derived directly from Equation 1:

$$\frac{du}{dt} = \frac{\alpha_1}{1 + v^{\beta}} - u, \quad \frac{dv}{dt} = \frac{\alpha_2}{1 + (u/(1 + [\text{IPTG}]/K_I)^{\mu})^{\gamma}} - v. \tag{2}$$

Note that some of the parameters from Equation 1 have been eliminated by rescaling protein concentrations and time into dimensionless forms. A similar model was constructed for the pIKE toggle. In the model of the pTAK toggle, the Lac repressor protein is synthesized at a rate that depends on the concentration of the λ repressor protein, whereas the λ repressor is synthesized at a rate that depends on concentration of LacI and on the second input to the inducible Lac switch, IPTG. The values of the relative maximal expression rates α_1 and α_2 are proportional to ν_{max} in Equation 1, and the coefficients β, γ, and μ determine the cooperativity of repressor-operator and repressor-inducer binding. The identical decay terms for the two proteins arise due to cell growth, which introduces a constant dilution of material within the cell. A similar phenomenological model was used in the construction of the ring oscillator (12).

Analysis of the toggle model gave valuable insights into rational design principles that were used to guide the experimental implementation of a working toggle switch. In particular, the model predicts that bistability is favored by high cooperativity (nonlinearity) of repressor binding (i.e., high values of β and γ) and comparable maximum expression rates α_1 and α_2. Figure 5*a* illustrates the approximate location of various pTAK and pIKE constructs in the α_1, α_2 parameter plane.

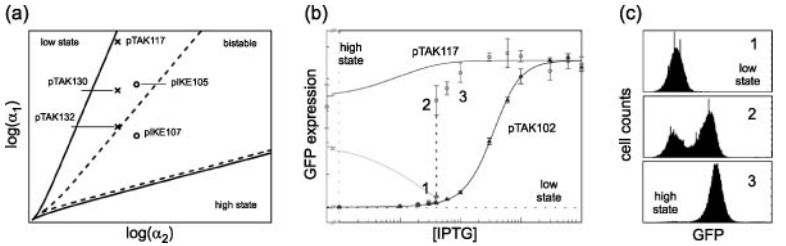

Figure 5 Toggle experiments. (*a*) The approximate location of various constructs in the α_1, α_2 (maximal expression rates) parameter plane. Full lines are for the LacI/λ repressor toggle (pTAK). Broken lines are for the LacI/TetR toggle (pIKE). (*b*) Fit of experimental data to the model. pTAK117 is a working toggle switch. pTAK102 is an inducible Lac switch obtained by removing the *cI* gene from pTAK117. (*c*) Histograms of GFP expression corresponding to the three points indicated in (*b*).

It also shows how modulation of α_1 can change the pIKE system from monostability (pIKE105) to bistability (pIKE107) and shift the pTAK system deeper into the bistable regime. While the toggle model lacks molecular detail and contains only a few adjustable parameters, it can nevertheless be fitted to experimental data. Figure 5*b* shows a model fit to the average population response, measured in terms of GFP fluorescence, to IPTG induction for both a bistable toggle construct (pTAK117) and an inducible Lac switch construct (pTAK102) obtained by eliminating the *cI* gene. While the model shows a good fit to the average population response, it cannot, by definition, capture the relatively wide range of responses of individual cells evident from the broad population distributions. The population distributions in Figure 5*c* were observed at different levels of IPTG. While the low and high expression states were indeed stable, the expression within individual cells may be significantly different from the average population expression level. Furthermore, as shown in Figure 5*c*, plot 2, a bimodal, double-peaked population distribution was observed at an intermediate level of IPTG. The bimodal distribution arises because, at the time of the measurement, some cells were in the low state and others in the high state. This differential response of individual cells cannot be captured by the deterministic model.

The deterministic model used to guide the construction of the toggle switch is an example of a top-down, phenomenological approach to the modeling of gene regulatory networks. Another approach is to build on the known molecular details using standard mass action kinetics. This approach is useful to explore what genetic modules are required and how they should be connected in order to achieve a network with desired properties. A number of networks based on the λ repressor control circuit have been explored theoretically using deterministic equations derived from mass action kinetics (17, 69, 73, 74). For example, mass action kinetics and the separation of timescales (the pseudo-steady-state approximation) can be employed to describe the λ repressor autoregulatory network,

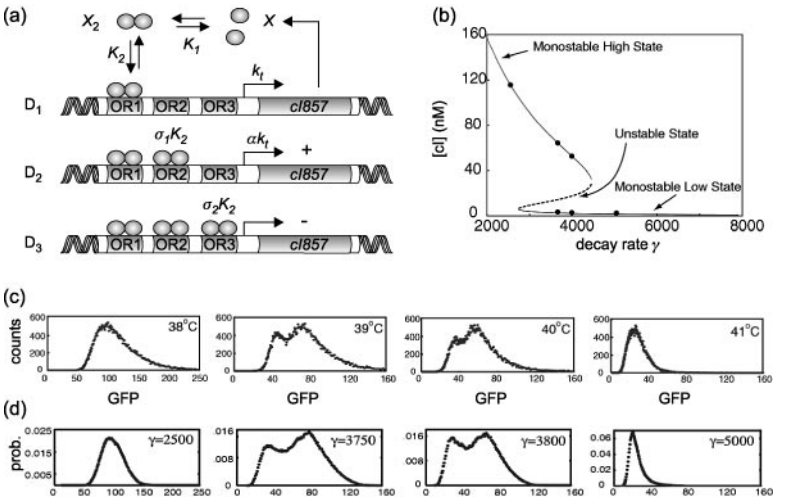

Figure 6 Model of λ repressor autoregulatory circuit. (*a*) Binding kinetics of λ repressor and definitions of promoter states. λ repressor monomers X combine into dimers X_2. The sequential binding of λ repressor to the OR1, OR2, and OR3 operators gives rise to three promoter states, D_1, D_2, and D_3. The equilibrium constants for these reactions are K_1, K_2, and $\sigma_1 K_2$, $\sigma_2 K_2$, respectively. Expression is low (rate k_t) from the promoter with no repressor bound (not shown) and the D_1 state and increased by a factor α from the D_2 state. Expression is absent from the D_3 state. (*b*) Steady-state diagram predicted by the deterministic model. High and low expression states coexist at intermediate values of the temperature-dependent decay (inactivation) rate γ. (*c*) Experimentally observed population distributions. The two peaks observed at 39°C and at 40°C indicate the coexistence of high and low expression states at these temperatures. (*d*) Population distributions obtained from a model that incorporates stochasticity in the expression of λ repressor (explained in Section 4.3).

illustrated schematically in Figure 6*a*, by a single ODE. In this model, dimerization of the temperature-sensitive λ repressor (denoted X) and the binding of λ repressor dimers (denoted X_2) to the operators OR1, OR2, and OR3 follow reversible binding kinetics with equilibrium constants K_1, K_2, and $\sigma_1 K_2$, $\sigma_2 K_2$, respectively. λ repressor dimers bind sequentially to the three operator sites, giving rise to three promoter states, denoted D_1, D_2, and D_3 in Figure 6, in addition to the unoccupied promoter. Because the association reactions follow second-order mass action kinetics, their rates depend on the cell volume, but this is often ignored because the growth and division of bacteria are quite complex [e.g., see (75–78)]. The volume dependence can often, but not always, be eliminated by using protein concentration (rather than the number of protein molecules) as the dynamic variable. When cells are assumed to obey an exponential growth law, the effect of volume increase is the emergence of a constant decay term in the ODE describing

the time-evolution of protein concentration as in the previously described toggle switch model.

The plot in Figure 6*b* illustrates one of the model predictions (79). The rate γ of λ repressor decay (inactivation), which can be adjusted by temperature modulation, determines whether the system is in a state of high or low expression from the P_{RM} promoter. Low values of γ (low temperature) should result in a single high state, whereas high values of γ (high temperature) should result in a single low state. At intermediate values, the system can be in either a low or a high expression state, which suggests the existence of a hysteresis loop at intermediate rates of λ repressor decay. When γ is increased from a low initial value, the model predicts that a population of cells will remain in the high state until γ increases beyond a value of roughly 4000. When γ is decreased from a high initial value, the model predicts that a population will remain in the low state until γ decreases below a value of roughly 3000.

Coexistence of high and low expression states was indeed observed when the network was implemented experimentally (79). The experiments were conducted in *E. coli* by inserting the P_{RM} promoter, controlling the expression of the *cI857* and *gfp* genes, into a high copy number plasmid. Figure 6*c* shows the population distributions of fluorescence obtained at four different temperatures. At 38°C, the distribution was broad and had a single peak at high fluorescence, corresponding to a high expression state. Two peaks were observed in the population distributions at 39°C and 40°C. These peaks correspond to cells that are in states of high and low expression. Finally, at 41°C, the distribution had a single peak at a relatively low fluorescence corresponding to a low expression state. A control experiment that used the wild-type (natural) *cI* gene did not show bimodality at otherwise identical conditions. Hence, as predicted by the model, there is a single high state when γ is low, a single low state when γ is high, and an intermediate range of γ where cells can be in either a high or a low expression state. However, hysteresis could not be observed experimentally. As discussed in more detail in Section 4.3, this is likely the result of high stochasticity in expression from the P_{RM} promoter.

4.2. Stochasticity in Gene Expression

The response of individuals within a population of genetically identical cells may be significantly different from the average population response [e.g., see (27, 80–86)]. The amount of protein present within single cells may vary widely across a population of identical cells in a homogeneous environment, as first measured after induction of the *lac* operon in *E. coli* (27, 84). While the contribution to cell individuality and population heterogeneity from probabilistic fluctuations in a small number of molecules has been discussed for many years [see, e.g., (27, 84, 87–89)], it was only recently that the origin of stochasticity, or noise, in gene expression was investigated using a methodology that directly couples theory and experiments (79, 80, 90, 91). These studies not only demonstrated how gene expression noise can be manipulated experimentally but also showed how relatively simple

stochastic models can capture the individuality of single cells. By their very nature, deterministic models can, at best, be used to predict (bottom-up approach) or to capture (top-down approach) the average behavior of a cell population or the time-averaged behavior of single cells, as discussed for both the toggle switch and the λ repressor autoregulation system in the previous section.

It is challenging to construct robust gene circuits when the individual logic modules produce broad output distributions with low signal-to-noise ratios. Therefore, from an engineering point of view, it is important to identify the "knobs" and "handles" that can be used to optimize network function, usually by minimizing the heterogeneity of the population response to an input signal. The level of heterogeneity depends on the mechanisms and kinetics of the individual logic modules as well as the architecture of the network. For instance, negative feedback has been shown to decrease noise (90, 92), whereas positive feedback is generally known to increase noise (93, 94). We are interested in constructing networks with all types of connectivity, and we focus therefore on the modulation of noise in isolated logic modules, i.e., in the transcription and translation of a single gene. Stochastic models have been used to study the effects of noise in more complex genetic networks (6, 36, 95–109) and we refer the reader to these papers for more in-depth studies.

A simple model of prokaryotic transcription and translation (91, 108) is illustrated schematically in Figure 7*a*. It involves probabilistic production of mRNA at an average rate κ_R; probabilistic translation of individual mRNA molecules into protein at an average rate κ_P; and probabilistic decay of mRNA and protein molecules at average rates γ_R and γ_P, respectively. All of these processes

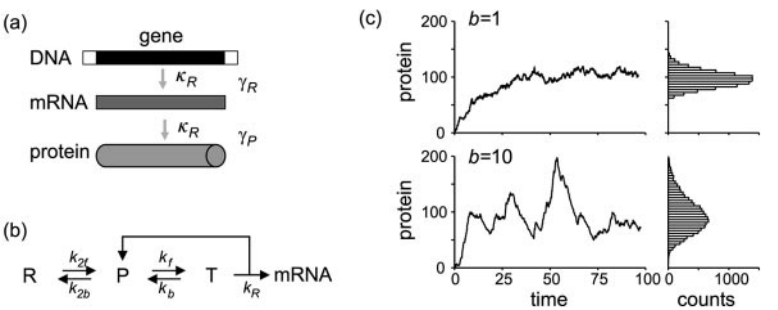

Figure 7 (*a*) Simple stochastic model of transcription and translation. Production of mRNA and protein from each mRNA occurs at average rates κ_R and κ_P, respectively. Decay of mRNA and of protein is pseudo–first order with rate constants γ_R and γ_P, respectively. (*b*) Simple model of transitions in a repressible prokaryotic promoter. P = unoccupied promoter, R = repressed promoter, T = promoter-RNA polymerase complex. The rate constants for association (k_{if}), dissociation (k_{ib}), and open complex formation (k_R) are pseudo–first order. (*c*) Illustration of the effect of increased burst parameter, $b = \kappa_P/\gamma_R$, on the evolution of protein number within single cells and on the steady-state population distribution.

are immensely complex and more detailed models have been investigated (16, 110, 111). Moreover, factors such as cell division and fluctuations in the number of RNA polymerases and ribosomes (and other factors involved in gene expression) may contribute significantly to cell-cell variability in protein content (111). These factors are difficult to manipulate without causing changes in cell function and are therefore not suitable to tune in the optimization of engineered gene regulatory networks. In contrast, the simple model illustrated in Figure 7 incorporates parameters that are available for direct experimental manipulation.

Figure 7*b* illustrates a more detailed model of a promoter region, which involves competitive binding of RNA polymerase and a repressor. In this model, binding of RNA polymerase to the unoccupied promoter (the P state) leads to the formation of a complex (the T state) that is able to initiate transcription at an average rate k_R. Binding of repressor leads to a promoter complex (the R state) that is unable to bind RNA polymerase. In addition, the model assumes that the concentrations of transcription factors and cofactors are constant, such that they can be absorbed into the pseudo-first-order rate constants for dissociation (k_{ib}) and association (k_{if}) of RNA polymerase ($i = 1$) and of repressor ($i = 2$). Models such as those illustrated in Figures 7*a* and 7*b* can be solved either numerically (112–115) or analytically using the so-called Master equation approach (101, 106, 111, 116–119).

When the timescale of RNA polymerase and repressor association/dissociation is fast relative to that of open complex formation (or to be more precise, when $k_{ib} \gg k_R$ for $k_{ib} < k_{if}$), the probability of transcription initiation is constant in time and the model in Figure 7*b* reduces to that in Figure 7*a*. In this case, the distribution of mRNA molecules within a population of cells, or measured in a single cell over a long period of time, follows a Poisson distribution with a variance σ_r^2 equal to the mean $\langle r \rangle$. Therefore, stochasticity in transcription contributes significantly to gene expression noise by converting a deterministic input, i.e., the average rate of transcription, into cell-cell variability in cellular mRNA content.

When the population distribution of mRNA is Poisson-distributed, it has been shown theoretically (91, 108) that the variance in protein content is approximately given by $\sigma_p^2 = \langle p \rangle (1 + b)$, where $b = \kappa_P / \gamma_R$ is the average number of proteins produced per mRNA transcript, which is referred to as the (protein) burst parameter. While the variance increases linearly with the rate of transcription, the ratio $\sigma_p^2 / \langle p \rangle$, a measure of noise called the Fano factor (91, 108) is independent of the mean number of proteins, $\langle p \rangle = \kappa_P \kappa_R / \gamma_R \gamma_P$. Another, and more common, measure of noise is the coefficient of variation ($\sigma_p^2 / \langle p \rangle^2$) and the two measures are equally meaningful. It is probably difficult to obtain population distributions of mRNA and protein content that are tighter than the Poisson distribution, and a variance that is equal to the mean, $\sigma_p^2 = \langle p \rangle$, i.e., a Fano factor of unity, may be considered at the basic "unit" of noise in gene regulatory networks. The Fano factor is thus a measure of noise that directly correlates the width of the population distribution to that of a Poisson distribution with the same mean. The coefficient of variation, which gives a direct measure of the signal-to-noise ratio, is also valuable, particularly when optimizing gene network function.

The validity of the predictions relating gene expression noise to the parameters associated with transcriptional and translational efficiency was demonstrated experimentally (91) by varying the expression of GFP from an IPTG-inducible promoter inserted into the chromosome of the prokaryote *Bacillus subtilis*. In one set of experiments, the rate of transcription was controlled by adding varying amounts of IPTG or by mutating the promoter region. In agreement with the model, the ratio $\sigma_P^2/\langle p\rangle$ was found to be relatively independent of the transcriptional efficiency, with a slightly increased level of noise at increased rates of transcription. This slight increase may be explained by a model where each cell has a slightly different value (drawn from a normal distribution) of the parameters (M. Kærn, unpublished results). In another set of experiments, the translational efficiency was manipulated by changing the DNA sequence of the ribosome binding site and the start codon of the *gfp* gene. In agreement with the model, the Fano factor was found to increase linearly with translational efficiency and the effect was much more pronounced than the effect observed at increased transcriptional efficiency.

The study described above provides direct experimental evidence for a rational design principle (110) that can be used to decrease cell-cell variability in an engineered gene regulatory network. Interestingly, this work showed that genes that are transcribed at high rates but inefficiently translated are less noisy than genes that are transcribed at low rates and efficiently translated. Hence, to obtain a tighter distribution with the same mean $\langle p\rangle$, the translational efficiency κ_P should be decreased and transcriptional efficiency κ_R increased while keeping their product $\kappa_P\kappa_R$ unchanged (91). This is illustrated in Figure 7*c*, which shows the evolution of the number of proteins in a single cell, as well as the steady-state population distribution (of 10^4 cells) for $\kappa_P\kappa_R = 10$, with two different values of the burst parameter. The two simulations give rise to the same average number of proteins, but the distribution is much broader in the simulation with the higher burst parameter.

Several factors that are not incorporated into the simple model in Figure 7*a* can affect the probability of producing a transcript and cause an increased cell-cell variability in mRNA content. Examples include slow association/dissociation of a transfactor from its binding site, referred to as operator fluctuations (101), and fluctuations in the amount (or activity) of transfactor molecules. Operator fluctuations are likely a significant source of transcriptional noise when the transfactor binds to, or is released from, its binding site with a relatively low probability (101). The effect of fluctuations in transfactor molecules was demonstrated in an experiment (80) designed to distinguish the contributions to overall gene expression noise from stochasticity in transcription and translation (referred to as intrinsic noise) from stochasticity in other factors (referred to as extrinsic noise). In this study, two identical LacI-repressed promoters expressing either CFP or YFP were inserted into the chromosome of *E. coli*, and the *lacI* gene was carried on a multicopy plasmid. The differences in CFP and YFP fluorescence in the same cell provided an estimate of stochasticity in the transcription and translation of the two genes (intrinsic noise), whereas differences in fluorescence of either CFP or YFP between cells gave an estimate of noise introduced by other factors, such as the

amount of repressor within individual cells (extrinsic noise). An increased level of extrinsic noise (measured in this study by the coefficient of variation) was observed at intermediate levels of IPTG induction. Based on careful model analysis (111), the effect was interpreted to arise from cell-cell variability in the number of repressor molecules (80). One might speculate that the absence of a similar effect when the *lacI* gene was integrated into the chromosome (91) indicates that the increased level of extrinsic noise arises from *lacI* being carried on a multicopy plasmid. It is known that plasmid replication and partitioning of plasmids at cell division introduces variability in the copy number of plasmid-borne genes [for a recent discussion, see (106)], which is likely to result in an increased cell-cell variability in protein content compared to the case where the gene is chromosomally carried.

4.3. Semiquantitative Modeling

The stochastic models discussed in the previous section are quite general and their predictions are applicable to gene expression in most prokaryotic cells and perhaps even to cells of higher organisms. It would be beneficial if quantitative models of specific genetic logic modules could be constructed. The TetR and LacI repressors are two of the best studied transcriptional regulators, and their structure and function are understood on the atomic level [see, e.g., (24–26, 31, 34)]. This wealth of information allows, in principle, for the formulation of detailed quantitative models of the logic modules described in Section 3.1. Unfortunately, quantitative molecular-level models involve a large number of undetermined parameters, which necessitates some degree of simplification. For example, the TetR dimer is known to bind two inducer molecules, leading to (at least) three different forms of the repressor molecule. Specifically, in the $P_{\mathrm{Ltet0-1}}$ promoter, the repressor can bind to two operators, and each operator can be unoccupied, occupied by dimer with no inducer, occupied by dimer with one inducer molecule, or occupied by dimer with two inducer molecules. This gives a total of 16 different promoter states and some 80 transitions, i.e., reactions, between these states. Rate constants, usually obtained in vitro, if at all, must be assigned to each of these transitions in order to simulate the inducible Tet switch based on the known interactions between the TetR repressor, inducer, and the *tet* operator.

The most pragmatic way to reduce the number of parameters while relying on the known reaction mechanisms (i.e., semiquantitative modeling), is to incorporate only the most essential states and assume that the transitions between different states follow generalized mass action kinetics. This approach allows known molecular mechanisms to be modeled explicitly with additional degrees of freedom that can be used to fit experimental data. The latter is necessary not only because of insufficient data regarding the underlying reaction mechanisms but also because of the crowded and imperfectly mixed microenvironment in which the reactions take place [e.g., see (71, 94, 120–122)]. Based on experimental and theoretical considerations regarding reaction order of elementary chemical reactions under dimensionality restricted conditions (for instance, one-dimensional diffusion of

proteins along the DNA), it has been suggested (19, 23) that the integer reaction orders derived from mass action kinetics should be replaced, where appropriate, with real-valued exponents (i.e., fractal kinetics).

A semiquantitative model of the inducible Tet switch could be constructed from the model illustrated in Figure 7*b* and the following reaction steps: (*a*) constitutive expression of TetR monomers, (*b*) dimerization of monomers, (*c*) simultaneous binding of two inducer molecules to one repressor dimer, and (*d*) simultaneous binding of two TetR dimers to *tet* operators. Using generalized mass action kinetics, the elementary reaction representation of steps (*c*) and (*d*) is given by

$$(\mathrm{TetR})_2 + \beta \mathrm{ATc} \overset{K_I}{\underset{\leftarrow}{\rightarrow}} (\mathrm{TetR} \cdot \mathrm{ATc})_2 \tag{3}$$

$$\gamma(\mathrm{TetR})_2 + P \overset{k'_{2f}}{\rightarrow} R, \tag{4}$$

where K_I is the (effective) equilibrium constant for ATc binding to TetR dimers and k'_{2f} is the (effective) rate constant for the binding of TetR dimers to the operators. The reaction orders β and γ can be determined experimentally or be estimated by assuming all-or-none binding in accordance with mass action kinetics, i.e., $\beta = 2$ and $\gamma = 2$.

When the dimerization and inducer binding reactions are fast relative to protein synthesis and decay, conservation of mass (detailed balance) can be used to relate the number of active TetR dimers to the total number of TetR dimers, n_{tet}. This number in turn determines the pseudo-first-order rate constant for repressor binding (the $P \rightarrow R$ transition in Figure 7*b*), which can be approximated by

$$k_{2f} = k'_{2f}(TetR)_2^{\gamma} = \frac{k'_{2f} n_{tet}^{\gamma}}{(1 + c_I^{\beta}[\mathrm{ATc}]^{\beta})^{\gamma}}, \tag{5}$$

where c_I is related to the equilibrium constant K_I. The population- or time-averaged rate of expression ($\langle\rangle$ denotes average) is given by $\langle \nu \rangle = k_R \langle T \rangle$, and the average relative rate of expression, denoted $\bar{\nu}$ for simplicity, can be obtained from the conservation of promoter copy number $m = \langle T \rangle + \langle P \rangle + \langle R \rangle$ as

$$\bar{\nu} = \frac{\langle \nu \rangle}{\langle \nu \rangle_{max}} = \frac{(1 + c_1)(1 + c_I^{\beta}[\mathrm{ATc}]^{\beta})^{\gamma}}{(1 + c_1)(1 + c_I^{\beta}[\mathrm{ATc}]^{\beta})^{\gamma} + c_2 n_{tet}^{\gamma}}, \tag{6}$$

where $c_1 = (k_R + k_{1b})/k_{1f}$, $c_2 = k'_{2f}/k_{2b}$, and $\langle \nu \rangle_{max}$ is the maximal rate of expression, $\langle \nu \rangle_{max} = m k_R c_1/(1 + c_1)$. In many cases, the value of $\bar{\nu}$ can be obtained simply by dividing the population-averaged fluorescence with that obtained under conditions of maximal expression. Most of the parameters in Equation 6 can be estimated by a fit to measurements of population-averaged expression, for instance, using the Hill-plot technique, i.e., by plotting $\ln \bar{\nu}/(1 - \bar{\nu})$ as a function of ATc or n_{tet} (or CFP expression in a cascading circuit).

While Equation 6 can be fitted to experimentally measured average population response, it cannot capture the response of individual cells. One approach to

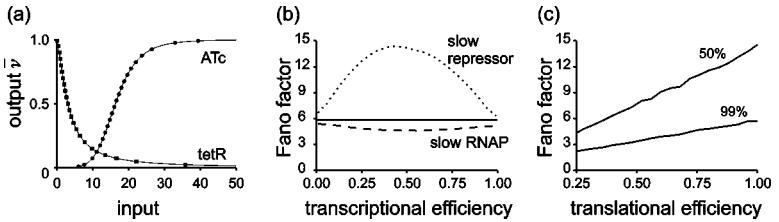

Figure 8 Stochastic simulations of the inducible Tet switch. (*a*) Population average response to modulation of repressor (*squares*) and inducer (*circles*). Full lines are obtained from Equation 6. (*b*) Dependence between expression noise and transcriptional efficiency (ATc induction) for constant probability of producing a transcript (*full line*) and slow dissociation of repressor (*dotted line*) and RNA polymerase (*broken line*). (*c*) Noise amplification by increased transcriptional efficiency for slow dissociation of repressor at two levels of induction.

capture this cell individuality is to model the stochasticity of transcription and translation as described in Section 4.2 and to fit the resulting population distributions to distributions obtained experimentally. Figure 8 shows the results of numerical simulations of the inducible Tet switch in a population of 10^4 cells. The outcome of each simulation is a population distribution similar to those shown in Figure 7*c*. The plot in Figure 8*a* shows the relationship between the inputs (ATc and n_{tet}) and the normalized output, $\overline{\nu}$, for simulations (*circles and squares*) and the deterministic model (*full curves*). The two methods predict the same average population response. Figures 8*b* and 8*c* illustrate how gene expression noise (the Fano factor) depends on transcriptional (modulation of ATc levels) and translational (modulation of κ_P values) efficiencies. The straight line in Figure 8*b* is obtained when the probability of producing a transcript is constant. The dotted and broken curves are obtained for slow dissociation of repressor and RNA polymerase, respectively, and illustrate how the level of gene expression noise may depend on the relative timescale of transcription factor kinetics [see also (101, 103, 111)]. Finally, the lines in Figure 8*c* illustrate how translation can amplify noise introduced by slow operator fluctuations at the level of transcription. At a high level of transcriptional noise (50% induction), the effects of increased translational efficiency are much more pronounced compared to a low level of transcriptional noise (99% induction) and the case where mRNA is Poisson-distributed (constant probability of producing a transcript). The simulation results in Figures 8*b* and 8*c* can be compared directly to experimental observations, and the model parameter can thus be fitted to capture the in vivo behavior of individual cells.

Simulation schemes that model each individual reaction event can become quite computationally demanding as the copy number of input and output molecules increases. An alternative and computationally efficient way to simulate a semi-quantitative stochastic model is to assume that the time-evolution of an input or output X, produced at rate $f(x)$ and degraded at rate $g(x)$, follows the stochastic

differential equation (SDE) given by

$$\frac{dx}{dt} = f(x) - g(x) + \sqrt{f(x) + g(x)}\xi_t, \tag{7}$$

where ξ_t is a stochastic process characterized by $\langle \xi_t \rangle = 0$ and $\langle \xi_t \xi_t' \rangle = \delta(t - t')$ (Gaussian white noise). An SDE-based model can be simulated using readily programmed algorithms (79, 118, 124) that run orders of magnitude faster than schemes where each reaction event is modeled explicitly.

The SDE approach treats the input/output signals as being continuous rather than discrete and cannot be employed in all situations [for a recent discussion, see (101)]. However, given appropriate conditions, it can be an extremely efficient and precise tool for the modeling of gene networks. This was recently demonstrated (79) using an SDE model to describe the λ repressor autoregulation experiments discussed in Section 4.1. In the study (79), it was assumed that the time-evolution of λ repressor monomers obeys Equation 7, where the deterministic description of λ repressor self-regulation (see Section 4.1) is augmented with a stochastic term. In addition, the model included an SDE equation describing GFP fluorescence and an equation describing cell growth. As it was illustrated in Figure 6*d*, the SDE model was able to reproduce the experimentally observed population distributions remarkably well. In agreement with experiments, bimodality of the population distribution was observed at intermediate temperatures, or rate of λ repressor inactivation, which is consistent with an underlying deterministic bistability when the noise amplitude is high enough to allow individual cells to make noise-induced transitions (117) between the low and the high expression states. In addition, the model was able to reproduce the parametric dependence of population variances quite well [see (79) for details].

5. CONCLUDING REMARKS

In this review, we have described an engineering approach to the construction of artificial gene regulatory networks and outlined a number of complementary approaches by which such networks can be modeled. Simple engineered gene network modules that allow controlled expression of one or two genes are commercially available and widely used in fundamental scientific, medical, and industrial research. The implementation of more complex networks, such as toggle switches, oscillators, cascading networks, and logic gates, has demonstrated sophisticated control of gene expression. It can be expected that these more complex gene network systems will be useful for multifaceted control of gene expression and of cellular behavior. Rather than being simple inverters and transistors, each of these networks possesses a specific computational logic, which, in the future, can be combined with other logic modules (constructed with different sets of input/output signals) to increase the computational power and functionality of an engineered gene regulatory network. Such networks may allow for predictable and robust control in complex and fluctuating cellular environments and thereby have a

significant impact on future biotechnologies. Engineered gene networks could, for example, be used to probe natural gene regulatory networks with multiple independent perturbations or act as sensors, monitoring multiple environmental factors and keeping track of past cellular and/or environmental events. Artificial gene networks that are constructed to mimic natural gene networks could provide a deeper understanding of how the latter are regulated and could possibly be used to replace or supplement natural cell function. In this context, it is worth mentioning the re-engineering of viral regulatory pathways, which has already shown promising results in cancer research [e.g., see (125–129)].

While simple gene networks can be constructed without the aid of mathematics, qualitative and quantitative models will become increasingly important as the number of network components and modules increases. We envision that appropriately detailed models can be used, for example, to identify the origins of network malfunction (i.e., for "debugging") or to suggest the optimal rational design strategy among redundant network architectures. Nonetheless, there are fundamental problems associated with the current modeling of gene regulation within the crowded and heterogeneous microenvironments of living cells. While attempts have been made to address some of these problems, new theoretical concepts and models are required to achieve a fundamental quantitative understanding of regulatory processes and organization within living cells [for a recent discussion see (71)]. It is our opinion, however, that the currently available modeling approaches, in careful combination with experiments, can be very powerful tools to construct, analyze, and interpret the function of individual genetic logic modules, as well as more complex gene circuits composed of these modules.

ACKNOWLEDGMENTS

We thank Diego DiBernado, Tim Gardner, Jeff Hasty, Farren Isaacs, and David McMillen for useful discussions and comments. This work was supported by DARPA and NSF.

The *Annual Review of Biomedical Engineering* is online at http://bioeng.annualreviews.org

LITERATURE CITED

1. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402:C47–52
2. Lauffenburger DA. 2000. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl. Acad. Sci. USA* 97:5031–33
3. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL. 2002. Hierarchical organization of modularity in metabolic networks. *Science* 297:1551–55
4. Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* 31:64–68
5. McAdams HH, Arkin A. 2000. Gene regulation: towards a circuit engineering discipline. *Curr. Biol.* 10:R318–20

6. McAdams HH, Shapiro L. 1995. Circuit simulation of genetic networks. *Science* 269:650–56
7. Gardner TS, Cantor CR, Collins JJ. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–42
8. Tchuraev RN, Stupak IV, Tropynina TS, Stupak EE. 2000. Epigenes: design and construction of new hereditary units. *FEBS Lett.* 486:200–2
9. Tropynina TS, Golubev OV, Stupak EE, Tchuraev RN. 2002. Construction of an artificial digenic network with epigenetic properties. *Mol. Biol.* 36:472–75
10. Guet CC, Elowitz MB, Hsing W, Leibler S. 2002. Combinatorial synthesis of genetic networks. *Science* 296:1466–70
11. Weiss R, Basu S. 2002. The device physics of cellular logic gates. In *NSC-1: The First Workshop on Non-Silicon Computing*. Boston, MA. pp. 54–61. Available from http://www.hpcaconf.org/hpca8
12. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–38
13. Endy D, Brent R. 2001. Modelling cellular behaviour. *Nature* 409:391–95
14. Arkin AP. 2001. Synthetic cell biology. *Curr. Opin. Biotechnol.* 12:638–44
15. De Jong H. 2002. Modeling and simulation of genetic regulatory systems: a literature review. *J. Comp. Biol.* 9:67–103
16. Gilman A, Arkin AP. 2002. Genetic "code": representations and dynamical models of genetic components and networks. *Annu. Rev. Genomics Hum. Genet.* 3:341–69
17. Hasty J, Isaacs F, Dolnik M, McMillen D, Collins JJ. 2001. Designer gene networks: towards fundamental cellular control. *Chaos* 11:207–20
18. Smolen P, Baxter DA, Byrne JH. 2000. Modeling transcriptional control in gene networks—methods, recent results, and future directions. *Bull. Math. Biol.* 62: 247–92
19. Savageau MA. 2001. Design principles for elementary gene circuits: elements, methods, and examples. *Chaos* 11:142–59
20. Rojo F. 2001. Mechanisms of transcriptional repression. *Curr. Opin. Microbiol.* 4:145–51
21. Jacob F, Monod J. 1961. Genetic regulatory mechanisms in synthesis of proteins. *J. Mol. Biol.* 3:318–56
22. Müller-Hill B. 1996. *The lac Operon*. New York: Walter de Gruyter
23. Müller-Hill B. 1998. Some repressors of bacterial transcription. *Curr. Opin. Microbiol.* 1:145–51
24. Bell CE, Lewis M. 2000. A closer view of the conformation of the Lac repressor bound to operator. *Nat. Struct. Biol.* 7:209–14
25. Friedman AM, Fischmann TO, Steitz TA. 1995. Crystal structure of *lac* repressor core tetramer and its implications for DNA looping. *Science* 268:1721–27
26. Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, et al. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271:1247–54
27. Novick A, Weiner M. 1957. Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci. USA* 43:553–66
28. Hillen W, Berens C. 1994. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annu. Rev. Microbiol.* 48:345–69
29. Kaneko M, Yamaguchi A, Sawai T. 1985. Energetics of tetracycline efflux system encoded by Tn10 in *Escherichia coli*. *FEBS Lett.* 193:194–98
30. Hillen W, Gatz C, Altschmied L, Schollmeier K, Meier I. 1983. Control of expression of the Tn10-encoded tetracycline resistance genes. Equilibrium and kinetic investigation of the regulatory reactions. *J. Mol. Biol.* 169:707–21
31. Hinrichs W, Kisker C, Duvel M, Muller A, Tovar K, et al. 1995. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* 264:418–20

32. Kedracka-Krok S, Wasylewski Z. 1999. Kinetics and equilibrium studies of Tet repressor-operator interaction. *J. Protein Chem.* 18:117–25
33. Kleinschmidt C, Tovar K, Hillen W, Porschke D. 1988. Dynamics of repressor-operator recognition: the Tn10-encoded tetracycline resistance control. *Biochemistry* 27:1094–104
34. Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. 2000. Structural basis of gene regulation by the tetatracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.* 7:215–19
35. Ptashne M. 1992. *A Genetic Switch. Phage λ and Higher Organisms*. Cambridge, MA: Cell Press & Blackwell Sci. Publ. 2nd ed.
36. Arkin A, Ross J, McAdams HH. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149:1633–48
37. Frankel AD, Kim PS. 1991. Modular structure of transcription factors: implications for gene regulation. *Cell* 65:717–19
38. Brent R, Ptashne M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43:729–36
39. Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1I2 regulatory elements. *Nucleic Acid Res.* 25:1203–10
40. Brent R, Ptashne M. 1984. A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene. *Nature* 312:612–15
41. Gossen M, Bonin AL, Bujard H. 1993. Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. *Trends Biochem. Sci.* 18:471–75
42. Blau HM, Rossi FMV. 1999. Tet B or not tet B: advances in tetracycline-inducible gene expression. *Proc. Natl. Acad. Sci. USA* 96:797–99
43. Belli G, Gari E, Piedrafita L, Aldea M, Herrero E. 1998. An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res.* 26:942–47
44. Brown M, Figge J, Hansen U, Wright C, Jeang KT, et al. 1987. *lac* repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in mammalian cells. *Cell* 49:603–12
45. David KM, Perrot-Rechenmann C. 2001. Characterization of a tobacco bright yellow 2 cell line expressing tetracycline repressor at a high level for strict regulation of transgene expression. *Plant Physiol.* 125:1548–53
46. Deuschle U, Pepperkok R, Wang F, Giordano TJ, McAllister WT, et al. 1989. Regulated expression of foreign genes in mammalian cells under the control of coliphage T3 RNA polymerase and lac repressor. *Proc. Natl. Acad. Sci. USA* 86:5400–4
47. Deuschle U, Hipskind RA, Bujard H. 1990. RNA polymerase II transcription blocked by *Escherichia coli* lac repressor. *Science* 248:480–83
48. Dingermann T, Frank-Stoll U, Werner H, Wissmann A, Hillen W, et al. 1992. RNA polymerase III catalysed transcription can be regulated in *Saccharomyces cerevisiae* by the bacterial tetracycline repressor-operator system. *EMBO J.* 11:1487–92
49. Figge J, Wright C, Collins CJ, Roberts TM, Livingston DM. 1988. Stringent regulation of stably integrated chloramphinicol acetyl transferase genes by *E. coli* lac repressor in monkey cells. *Cell* 52:713–22
50. Fuerst TR, Fernandez MP, Moss B. 1989. Transfer of the inducible lac repressor/operator system from *Escherichia coli* to a vaccinia virus expression vector. *Proc. Natl. Acad. Sci. USA* 86:2549–53
51. Gatz C, Frohberg C, Wendenburg R. 1992. Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. *Plant J.* 2:397–404
52. Gatz C, Quail PH. 1988. Tn10-encoded

tet repressor can regulate an operator-containing plant promoter. *Proc. Natl. Acad. Sci. USA* 85:1394–97
53. Gossen M, Bujard H. 1992. Tight control of gene activity in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* 89:5547–51
54. Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, Bujard H. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* 268:1766–69
55. Hu MC, Davidson N. 1987. The inducible Lac operator-repressor system is functional in mammalian cells. *Cell* 48:555–66
56. Luukkonen BG, Seraphin B. 1998. Construction of an in vivo-regulated U6sn-RNA transcription unit as a tool to study U6 function. *RNA* 4:231–38
57. Nagahashi S, Nakayama H, Hamada K, Yang H, Arisawa M, Kitada K. 1997. Regulation by tetracycline of gene expression in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 255:372–75
58. Ulmasov B, Capone J, Folk W. 1997. Regulated expression of plant tRNA genes by the prokaryotic Tet and Lac repressors. *Plant Mol. Biol.* 35:417–24
59. Struhl K. 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* 98:1–4
60. Ptashne M, Gann A. 1997. Transcriptional activation by recruitment. *Nature* 386:569–77
61. Fussenegger M, Morris RP, Fux C, Rimann M, von Stockar B, et al. 2000. Streptogramin-based gene regulation systems for mammalian cells. *Nat. Biotechnol.* 18:1203–8
62. Weber W, Fux C, Daoud-el Baba M, Keller B, Weber CC, et al. 2002. Macrolide-based transgene control in mammalian cells and mice. *Nat. Biotechnol.* 20:901–7
63. Shimizu-Sato S, Huq E, Tepperman JM, Quail PH. 2002. A light-switchable gene promoter system. *Nat. Biotechnol.* 20:1041–44
64. Weber W, Fussenegger M. 2002. Artificial mammalian gene regulation networks—novel approaches for gene therapy and bioengineering. *J. Biotechnol.* 98:161–87
65. Villaverde A, Benito A, Viaplana E, Cubarsi R. 1993. Fine regulation of *cI857*-controlled gene expression in continuous culture of recombinant *Escherichia coli* by temperature. *Appl. Environ. Microbiol.* 59:3485–87
66. Barkai N, Leibler S. 2000. Circadian clocks limited by noise. *Nature* 403:267–68
67. Vilar JM, Kueh HY, Barkai N, Leibler S. 2002. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA* 99:5988–92
68. Weiss R, Knight T. 2000. Engineered communications for microbial robotics. In *DNA6: Sixth International Meeting on DNA Based Computers, DNA 2000*. Boston: Springer-Verlag
69. McMillen D, Kopell N, Hasty J, Collins JJ. 2002. Synchronizing genetic relaxation oscillators by intercell signaling. *Proc. Natl. Acad. Sci. USA* 99:679–84
70. Aubrecht J, Manivasakam P, Schiestl RH. 1996. Controlled gene expression in mammalian cells via a regulatory cascade involving the tetracycline transactivator and *lac* repressor. *Gene* 172:227–31
71. Kuthan H. 2001. Self-organisation and orderly processes by individual protein complexes in the bacterial cell. *Prog. Biophys. Mol. Biol.* 75:1–17
72. Halling PJ. 1989. Do the laws of chemistry apply to living cells? *Trends Biochem. Sci.* 14:317–18
73. Hasty J, Dolnik M, Rottschafer V, Collins JJ. 2002. Synthetic gene networks for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* 88:148101
74. Hasty J, Pradines J, Dolnik M, Collins JJ. 2000. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. USA* 97:2075–80
75. Keasling JD, Kuo H, Vahanian G. 1995. A Monte Carlo simulation of the

Escherichia coli cell cycle. *J. Theor. Biol.* 176:411–30
76. Koch AL. 1999. The re-incarnation, re-interpretation and re-demise of the transition probability model. *J. Biotechnol.* 71:143–56
77. Koppes LJ, Woldringh CL, Grover NB. 1987. Predicted steady-state cell size distributions for various growth models. *J. Theor. Biol.* 129:325–35
78. Tyson JJ, Diekmann O. 1986. Sloppy size control of the cell division cycle. *J. Theor. Biol.* 118:405–26
79. Isaacs FJ, Hasty J, Cantor CR, Collins JJ. 2003. Prediction and measurement of an autoregulatory genetic module. In press
80. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* 297:1183–86
81. Fiering S, Northrop JP, Nolan GP, Mattila PS, Crabtree GR, Herzenberg LA. 1990. Single cell assay of a transcription factor reveals a threshold in transcription activated by signals emanating from the T-cell antigen receptor. *Genes Dev.* 4:1823–34
82. Ko MSH, Nakauchi H, Takahashi N. 1990. The dose dependence of glucocorticoid-inducible gene expression results from changes in the number of transcriptionally active templates. *EMBO J.* 9:2835–42
83. Levin MD, Morton-Firth CJ, Abouhamad WN, Bourret RB, Bray D. 1998. Origins of individual swimming behavior in bacteria. *Biophys. J.* 74:175–81
84. Maloney PC, Rotman B. 1973. Distribution of suboptimally induced β-d-galactosidase in *Escherichia coli*. *J. Mol. Biol.* 73:77–91
85. Spudich JL, Koshland DR Jr. 1976. Non-genetic individuality: chance in the single cell. *Nature* 262:467–71
86. Sumner ER, Avery SV. 2002. Phenotypic heterogeneity: differential stress resistance among individual cells of the yeast *Saccharomyces cerevisiae*. *Microbiology* 148:345–51
87. Delbrück M. 1945. The burst size distribution in the growth of bacterial viruses (bacteriophages). *J. Bacteriol.* 50:131–35
88. Powell EO. 1958. An outline of the pattern of bacterial generation times. *J. Gen. Microbiol.* 18:382–417
89. Rahn O. 1932. A chemical explanation of the variability of the growth rate. *J. Gen. Physiol.* 15:257–77
90. Becskei A, Serrano L. 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590–93
91. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31:69–73
92. Savageau MA. 1974. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* 252: 546–49
93. Delbrück M. 1940. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* 8:120–24
94. Epstein IR. 1995. The consequences of imperfect mixing in autocatalytic chemical and biological systems. *Nature* 374: 321–27
95. Berg OG, Paulsson J, Ehrenberg M. 2000. Fluctuations and quality of control in biological cells: zero-order ultrasensitivity reinvestigated. *Proc. Natl. Acad. Sci. USA* 79:1228–36
96. Cook DL, Gerber AN, Tapscott SJ. 1998. Modeling stochastic gene expression: implications for haploinsufficiency. *Proc. Natl. Acad. Sci. USA* 95:15641–46
97. Endy D, Kong D, Yin J. 1997. Intracellular kinetics of a growing virus: a genetically structured simulation for bacteriophage T7. *Biotechnol. Bioeng.* 55:375–89
98. Endy D, You LC, Yin J, Molineux IJ. 2000. Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. *Proc. Natl. Acad. Sci. USA* 97: 5375–80
99. Gonze D, Halloy J, Gaspard P. 2002. Biochemical clocks and molecular noise:

theoretical study of robustness factors. *J. Chem. Phys.* 116:10997–1010
100. Kastner J, Solomon J, Fraser S. 2002. Modeling a hox gene network *in silico* using a stochastic simulation algorithm. *Dev. Biol.* 246:122–31
101. Kepler TB, Elston TC. 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* 81:3116–36
102. Kierzek AM, Zaim J, Zielenkiewicz P. 2001. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J. Biol. Chem.* 276:8165–72
103. Ko MSH. 1991. A stochastic model for gene induction. *J. Theor. Biol.* 153:181–94
104. Morton-Firth CJ, Bray D. 1998. Predicting temporal fluctuations in an intracellular signalling pathway. *J. Theor. Biol.* 192:117–28
105. Paulsson J, Berg OG, Ehrenberg M. 2000. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA* 97: 7148–53
106. Paulsson J, Ehrenberg M. 2001. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* 34:1–59
107. Roy S, Bose I, Manna SS. 2001. A cooperative stochastic model of gene expression. *Int. J. Mod. Phys. C* 12:413–20
108. Thattai M, van Oudenaarden A. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* 98:8614–19
109. Thattai M, van Oudenaarden A. 2002. Attenuation of noise in ultrasensitive signaling cascades. *Biophys. J.* 82:2943–50
110. McAdams HH, Arkin A. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* 94:814–19
111. Swain PS, Elowitz MB, Siggia ED. 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* 99:12795–800
112. Gibson MA, Bruck J. 2000. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A* 104:1876–89
113. Gillespie DT. 2000. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* 4:1716–33
114. Haseltine EL, Rawlings JB. 2002. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.* 117:6959–69
115. Kierzek AM. 2002. STOCKS: STOChastic Kinetic Simulations of biochemical systems with gillespie algorithm. *Bioinformatics* 18:470–81
116. Gardiner CW. 1990. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Berlin: Springer-Verlag
117. Horsthemke W, Lefever R. 1984. *Noise Induced Transitions. Theory and Applications in Physics, Chemistry and Biology*. Berlin: Springer-Verlag
118. Kloeden PE, Platen E. 1992. *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer-Verlag
119. van Kampen NG. 1992. *Stochastic Processes in Chemistry and Physics*. Amsterdam: North-Holland
120. Hess B, Mikhailov A. 1995. Microscopic self-organization in living cells: a study of time-matching. *J. Theor. Biol.* 176:181–84
121. Metzler R. 2001. The future is noisy: the role of spatial fluctuations in genetic switching. *Phys. Rev. Lett.* 87:068103
122. Smolen P, Baxter DA, Byrne JH. 1999. Effects of macromolecular transport and stochastic fluctuations on dynamics of genetic regulatory systems. *Am. J. Physiol.* 277:C777–90
123. Savageau MA. 1995. Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *J. Theor. Biol.* 176:115–24
124. Honeycutt RL. 1992. Stochastic Runge-Kutta algorithms. I. White noise. *Phys. Rev. A* 45:600–10

125. Dubensky TW. 2002. (Re-)engineering tumor cell-selective replicating adenoviruses: a step in the right direction toward systemic therapy for metastatic disease. *Cancer Cell* 1:307–9
126. Johnson L, Shen A, Boyle J, Kunich L, Pandey K, et al. 2002. Selectively replicating adenoviruses targeting deregulated E2F activity are potent, systemic antitumor agents. *Cancer Cell* 1:325–37
127. Nettelbeck DM, Rivera AA, Balague C, Alemany R, Curiel DT. 2002. Novel oncolytic adenoviruses targeted to melanoma: specific viral replication and cytolysis by expression of E1A mutants from the tyrosinase enhancer/promoter. *Cancer Res.* 62:4663–70
128. Ramachandra M, Rahman A, Zou AH, Vaillancourt M, Howe JA, et al. 2001. Re-engineering adenovirus regulatory pathways to enhance oncolytic specificity and efficacy. *Nat. Biotechnol.* 19:1035–41
129. Vile RG. 2001. Vironcology—not yet, but soon? *Nat. Biotechnol.* 19:1020–22