Annu. Rev. Biomed. Eng. 2001. 3:391–419

# CONTROL MOTIFS FOR INTRACELLULAR REGULATORY NETWORKS*

Christopher V. Rao[1] and Adam P. Arkin[2]
[1]*Department of Bioengineering, University of California, Berkeley, CA 94720; e-mail: c_rao@lbl.gov*
[2]*Departments of Bioengineering and Chemistry, University of California, Berkeley Physical Biosciences Division, Lawrence Berkeley National Laboratory, Howard Hughes Medical Institute, Berkeley, CA 94720; e-mail: aparkin@lbl.gov*

**Key Words** genetic circuits, biochemical reaction networks, genetic regulation, signal transduction, biological control

■ **Abstract** A number of technological innovations are yielding unprecedented data on the networks of biochemical, genetic, and biophysical reactions that underlie cellular behavior and failure. These networks are composed of hundreds to thousands of chemical species and structures, interacting via nonlinear and possibly stochastic physical processes. A central goal of modern biology is to optimally use the data on these networks to understand how their design leads to the observed cellular behaviors and failures. Ultimately, this knowledge should enable cellular engineers to redesign cellular processes to meet industrial needs (such as optimal natural product synthesis), aid in choosing the most effective targets for pharmaceuticals, and tailor treatment for individual genotypes. The size and complexity of these networks and the inevitable lack of complete data, however, makes reaching these goals extremely difficult. If it proves possible to modularize these networks into functional subnetworks, then these smaller networks may be amenable to direct analysis and might serve as regulatory motifs. These motifs, recurring elements of control, may help to deduce the structure and function of partially known networks and form the basis for fulfilling the goals described above. A number of approaches to identifying and analyzing control motifs in intracellular networks are reviewed.

## CONTENTS

## INTRODUCTION

One of the major challenges in the "postgenomic era" will be to deduce the behavior of genetic and biochemical networks. In order to grapple with their size and complexity, we need a conceptual framework to organize and analyze these systems. A challenge is that, unlike most physical systems, there are often too many ambiguities in biological systems for direct analytical or numerical analysis. One possible strategy is to use tools and concepts from the engineering sciences, control and systems theory in particular. Unlike in physical systems, we can invoke teleological arguments in biology: Networks have evolved to perform a physiological function. In this regard, engineered and biological systems share many common properties. A significant degree of the complexity in biological networks, however, cannot be attributed solely to the function itself, but rather to ensuring that the network performs robustly despite uncertain environments (cf. (1)). Any successful framework for analyzing biological networks, therefore, needs to address how systems perform robustly in an uncertain and dynamic environment. Furthermore, system failure, for both biological and engineered systems, can arise either from explicit loss in function, leading to total failure, or from a breakdown in the regulatory system, leading to more erratic failure observed in many genetic diseases (e.g. (2)). Thus, one needs to focus on the regulatory properties of the network.

In order to facilitate a functional and control approach for analyzing biological networks, many researchers have recently advocated a modular approach (3–5). A modular decomposition provides a convenient abstraction for deducing the behavior of complex networks (6, 7). A module is defined loosely as a subsystem whose function is separable from the function of other modules. One of the central challenges in biological control analysis will be determining the best ways to modularize a particular network or process. Consider, for example, gene expression in prokaryotes. We can view gene expression as a module, because its function is discernible from other physiological processes. Simultaneously, we can further

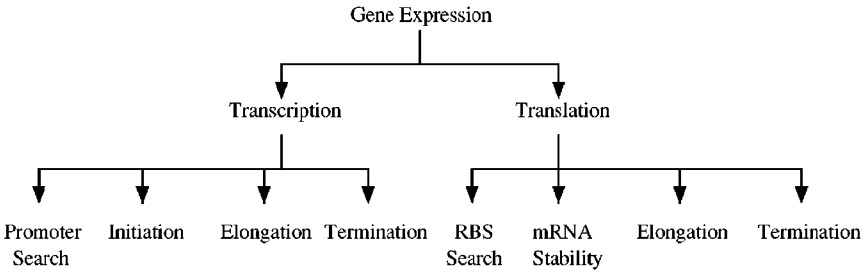



**Figure 1** A hierarchical decomposition of gene expression in prokaryotes.

decompose gene expression into a subset of separable functional modules (see Figure 1), thus illustrating the hierarchical organization of a modular decomposition. Likewise, we can view abstract gene expression as a parameterizable module that is used in multiple networks. This module is instantiated frequently in different pathways (e.g. *lac* and arginine biosynthesis). These instantiations may be more or less similar to each other. For example, *lac* and *ara* have similar regulatory mechanisms. At each level of the modular hierarchy, different mathematical models and different data may apply. Modules are now defined rather haphazardly. They may be defined as pathways such as glycolysis or the pentose phosphate cycle, by a class of functions such as "signal transduction" or "metabolism," by stoichiometric structure such as a "moiety-conserved" cycle, or by structural class such as "G-protein coupled signal transduction," etc. There are also more mathematical definitions of modules that define whether or not mass/energy flows between modules or only information. These are complementary ways of decomposing networks and each proves useful in different situations. However, a more formal approach for classifying subnetworks is in order.

Well characterized functional modules exist for a number of biological systems. We can begin to understand the principles governing intracellular regulation by understanding the different regulatory mechanisms used by functional modules. By analyzing the regulatory mechanisms in different functional modules, we may distill and catalogue recurring themes in regulation. We call these recurring themes "regulatory motifs." These motifs include both the biochemical mechanisms and the network topology used in control. Examples of regulatory motifs include autoregulation in gene expression and feedback inhibition in metabolism. These motifs should facilitate our investigations by uncovering some of the design principles involved in intracellular regulation. We may speculate, furthermore, that conserved motifs may have been selected for by evolution (cf. (8)). We may also begin to ask questions such as: what does this motif accomplish, why has this motif evolved rather than other motifs, why is the network more complex than it apparently needs to be, and how do different organisms instantiate a common motif for a given physiological process? The goal is to develop insight into how intracellular networks operate and are regulated.

Implicit is the use of computer simulation and mathematical analysis to elucidate the behavior of biological networks. The complexity of these networks and the increasing availability of data have led to a resurgence of interest in applying mathematical and engineering analysis to biological problems. From the dawn of modern systems theory, researchers have recognized the potential of applying engineering analysis to biology (9, 10). Whereas in the past, there were few data and fewer characterized systems amenable to analysis, today we find ourselves with the opposite problem. With the recent explosion in high-throughput technologies such as high-density oligonucleotide arrays for exploring the genetic regulatory architecture of the cell (cf. (11–14)) and the availability of extensive databases cataloging genes, proteins, and metabolic pathways (cf. (15–18)), we are no longer limited by the availability of experimental data. The challenge, rather, is to interpret and integrate all of the data available. The volume and complexity of the data necessitate computational and mathematical analysis (19). Whether these data are sufficient to deduce the behavior and regulatory properties of cellular physiology remains an open question.

The aim of this chapter is to review the current means by which intracellular networks are analyzed. Our review is not comprehensive and highlights only a fraction of the exciting work in this nascent field. For an alternative perspective, the reader is directed to the review articles (3, 20–24). Our review is organized as follows. In order to demonstrate what is meant by a regulatory motif, we first present a series of regulatory models, each representing a functional module or different physiological process. Each model describes a different example of a regulatory motif from the perspective of a control or systems engineer. These are all examples of "reverse engineering" wherein the function and structure of a network is deduced from data. Forward engineering methods, wherein a network is designed to meet certain specifications, also provides a route to understanding how to implement control strategies in biological media. In all cases three questions are addressed: (*a*) What is the control process that the cell is expressing? (*b*) How is this control implemented by the network of molecular interactions? and (*c*) How can the same control motif be instantiated using different molecular networks? Indeed, we seek to focus on the regulatory properties instead of the functional properties of the network. We conclude by offering our perspective on some of the challenges to forward and reverse engineering biological systems.

## MODELS OF CELLULAR REGULATION

Control and dynamic analyses of a number of central processes are outlined in this section. Examples are chosen both because they are rather well-worked out systems and because they are each representative of some canonical control task that might be amenable to systems analysis. These examples are: (1) metabolism, an example of continuous, deterministic control largely directed to maintenance of homeostasis, (2) bacterial chemotactic signal transduction as an example of a

signal detection, amplification, and tracking problem, (3) type-1 phase variation in *Escherichia coli* and the phage λ lysis/lysogeny decision as two examples of genetic switches, (4) T7 growth dynamics as an example of a linear manufacture control process, (5) the cell cycle as an example of a periodic scheduling process, and (6) gene expression networks as an example of how different levels of abstraction can be used to analyze network dynamics. This is far from a complete list of interesting control phenomena and their possible physical implementations. However, we hope they may serve as a starting point for others interested in biological systems analysis.

## Metabolism

Metabolism is analogous in many ways to a chemical plant or an oil refinery (25). In both systems, there are a large number of products simultaneously being decomposed and synthesized. Both processes are highly integrated and use energy derived from certain reactions in order to drive other reactions. Also, the concept of flux in metabolism is equivalent to flow rates in a chemical process. Both processes use a complex series of nested and cascaded feedback loops to ensure process flexibility and resiliency in the face of environmental changes and demands. In metabolism, feedback inhibits the excessive buildup of intermediate metabolites and maintains balanced growth. Similarly in chemical processes, feedback prevents intermediate operations from overloading and maintains an optimal and balanced distribution of products. In both systems, the dominant regulatory motif is negative feedback by end-product inhibition. Furthermore, many of the questions we ask in chemical process control can also be directed toward metabolic regulation. Several common questions arise in chemical process analysis (26): (*a*) Do feedback loops interact with other feedback loops? (*b*) How are feedback loops decoupled? (*c*) How are feedback loops integrated? and (*d*) What are the effects of recycling metabolites and enzymes?

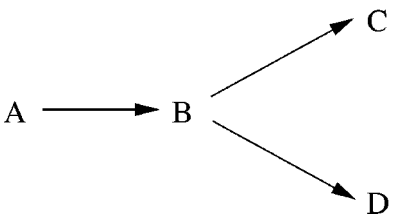
Understanding the mechanisms involved in metabolic regulation has important implications both in biotechnology and in medicine. For example, any rational attempt to engineer metabolism, say to redirect metabolic flux for the production of key metabolites or for enzyme-replacement gene therapy, requires at least an elementary understanding of how metabolism is regulated (27, 28). Increasing the expression of a single "rate-limiting" enzyme or inhibiting the flux through a particular branch does not always yield the desired change in flux, as the network is resilient to local changes (29, 30). Redirecting flux often requires a directed change in the activity of many enzymes (cf. (31)). A goal in unraveling the mechanisms in metabolic regulation is to understand the effect of an individual enzyme or metabolite on the global properties of the integrated pathway.

Many of the reactions constituting metabolism have been characterized for a number of organisms (cf. (32)), and kinetic models have been developed for many different aspects of metabolic regulation (for a review see (23, 33–35)). Examples include the glycolysis/gluconeogenesis switch (36), red blood cell metabolism (37–40), and metabolism in *Mycoplasma genitalium* (41). Still, little is known as

to how these integrated pathways are regulated. The complexity of these pathways has led to the development of numerous strategies for investigating the distribution and the control of flux in metabolism.

The reactions constituting metabolism in a particular cell form a complex network through which flux is distributed. One can characterize all possible distributions by analyzing the mass-conversion stoichiometry of the overall reaction network. This approach is known as flux balance analysis. The stoichiometry defines geometrically a nonnegative convex cone characterizing the feasible pathways through the network. One can explore the geometry of this convex cone using linear programming and speculate as to how flux is distributed for optimal growth (e.g. (42)) and redistributed to maintain balanced growth when certain pathways are eliminated (e.g. (43, 44)). This analysis, in some ways, resembles network flows in systems engineering (45). However, the extent of this analysis is limited when applied to the regulation of metabolism, as illustrated in the following example.

Consider the following branched pathway:




The flux in this pathway is distributed between the pathways $B \rightarrow C$ and $B \rightarrow D$. If the product $C$ is necessary for growth, then we expect flux through pathway $B \rightarrow C$. If we eliminate pathway $B \rightarrow D$, then all of the flux is directed through pathway $B \rightarrow C$. That is the extent of stoichiometric analysis, though how to maximize flux for larger, more constrained systems is more complex. Nowhere is regulation considered, even though it can profoundly affect the distribution and redirection of flux (29). For example, suppose the metabolite $D$ positively regulates an enzyme in pathway $A \rightarrow B$; then by eliminating pathway $B \rightarrow D$ one may actually reduce the flux through pathway $B \rightarrow C$. Though trivial, this example demonstrates a limitation of flux balance analysis. One also can use stoichiometry to decompose a network into a set of elementary flux modes or extreme currents (cf. (46–49)). These elementary modes define a unique set of independent routes through the metabolic network. In the example above, the two elementary modes are $A \rightarrow C$ and $A \rightarrow D$. It is not self-evident what role elementary modes have in metabolic regulation. Are these the "controllable" pathways of the network? Although stoichiometry alone provides little insight into the regulatory mechanisms in metabolism, it plays an important role in metabolic control analysis and the structural stability of reaction networks (cf. (46, 50, 51)).

Metabolic control analysis (MCA) (52, 53) measures the quantitative effect of an individual enzyme on the pathway flux and metabolite concentrations. The central concept in MCA is the control coefficient, a first-order measure of the change

in steady-state flux relative to a change in the reaction activity. There is also the related concept of an elasticity coefficient, a first-order measure of how an isolated enzyme or metabolite affects the reaction activity. The control coefficients describe the global properties of enzymatic activity, whereas the elasticity coefficients describe the local properties of enzymatic activity. Two theoretical results known as the summation and connectivity theorems relate the control coefficients (global network properties) to the elasticity coefficients (local network properties), thus allowing one to infer the effect of enzymatic activity on the overall pathway solely from its isolated effect on each reaction. One limitation of MCA is that it provides information only locally around a stationary regime, though strategies exist that extend its range. There is a large body of research in MCA that we do not attempt to survey. The reader is directed to the articles (54, 55) and the books (56, 57) for an overview of MCA. Many of the concepts of MCA are similar to classic control theory. By defining a transfer function for each reaction, analogous to the elasticity coefficient in MCA, one can determine the closed-loop properties of the system and, indirectly, the control coefficients. We note finally that the use of isotope labeling and nuclear magnetic resonance (NMR) for quantifying flux in vivo (cf. (58–61)) coupled with flux balance analysis and MCA offers tremendous potential for analyzing the regulatory properties of metabolism outside of large dynamic changes.

## Signal Transduction: Bacterial Chemotaxis

Chemotaxis is the process by which cells move and respond to changes in their environment (62) and, thus, is an example of signal transduction. From an engineering perspective, chemotaxis is a biological example of guidance control and signal amplification. Furthermore, microbial chemotaxis demonstrates how bacteria accommodate physical limitations associated with their size. In particular, bacteria are unable to sense concentration gradients spatially owing to their small size and can only sense temporal changes (63). Bacterial chemotaxis, therefore, requires a rudimentary form of memory (64). Bacteria also sense gradients over a wide range of concentrations, far in excess of what is expected from their limited number of receptors. A related feature of chemotaxis is adaptation: Bacteria respond only to changes in their environment rather than to the absolute state of their environment (65). Bacteria are also incapable of precise motion. Motion, rather, is analogous to a biased random walk. Bacteria control their motion by alternating between runs, where the cell moves in a relatively straight direction, and tumbles, where the cell randomly changes direction. Chemotaxis, therefore, possesses a number of interesting control issues.

Figure 2 is a rough schematic of the basic intracellular network controlling chemotaxis in *E. coli* (cf. (66–68)). Numerous models of chemotaxis have been proposed over the years (69–75). One of the main difficulties in modeling chemotaxis is providing a mechanism for adaptation, namely, the restoration of prestimulus behavior in the presence of attractant or repellent. The first two models to achieve a mechanistic description of adaptation were proposed by Hauri & Ross

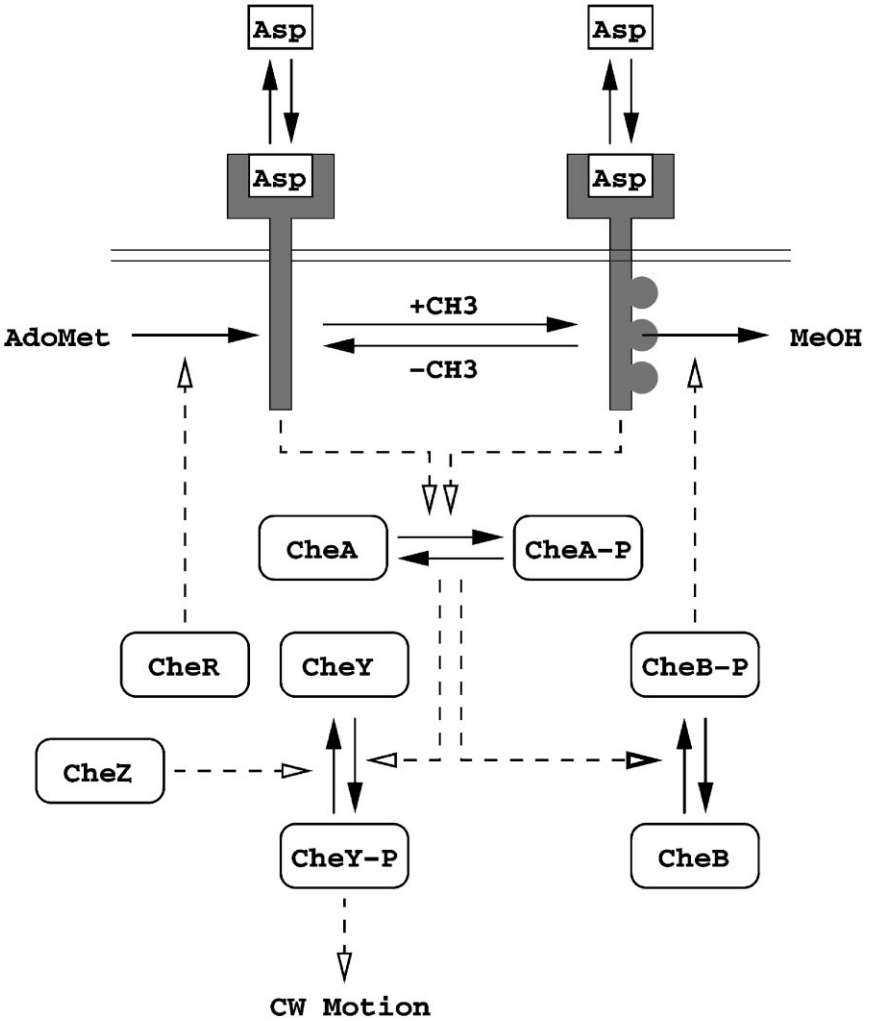



**Figure 2** The chemotaxis signal transduction pathway for aspartate response in *E. coli*. When aspartate binds to the receptor, its activity decreases and the corresponding rate of CheA autophosphorylation also decreases. Consequently, the rate CheY and CheB phosphorylation decrease. Decreased levels of CheY-P increase the likelihood of counter-clockwise rotation (run), as the binding of CheY-P to motor induces clockwise rotation (tumble). Decreased levels of CheB-P cause an increase in receptor methylation, as CheB-P removes the methyl groups from the receptor. Activity increases with the degree of receptor methylation and eventually returns to the prestimulus value, thus yielding adaptation. In the full system, there are different stimulants and a plethora of receptors. These signals must be integrated to steer the cell.

(72) and Spiro and coworkers (73). To achieve adaptation, the parameters in both models had to be finely tuned. Barkai & Leibler subsequently proposed a model, which was later expanded by Morton-Firth and coworkers (75), where exact adaptation does not depend on the parameterization of the network, but rather appears to be a robust property of the network. Using an elementary model of the methylation dynamics, Barkai & Leibler demonstrate that adaptation holds over a wide range of parameter values. In a subsequent paper (76), Leibler and coworkers experimentally validated the central hypothesis of their model: that adaptation is a robust property of the network. We note that although adaptation is a robust property of the network, chemotaxis may not be. The response time varies widely as a function of the kinetic parameters and protein concentrations, thus suggesting that the bacteria may be unable to find food sources or avoid repellents because they overrun their target or because they simply respond too slowly. Further work is needed to address this issue.

How does the network confer robustness? The model proposed by Barkai & Leibler demonstrated the effect but did not put it in an engineering context. Yi and coworkers (77) analyzed the chemotaxis network and demonstrated that it possesses an integrator. Their analysis was motivated by the "internal model principle" of control theory (78, 79), which states that a controller must contain a model of any disturbance it rejects or signal it tracks. One interpretation is that in order to adaptively respond to one's environment, one has to anticipate all possible contingencies (the internal model). For bacterial chemotaxis, a model in the form of an integrator is necessary for adaptation, where the contingencies are different concentrations of attractant and repellent. Loosely speaking, the integrator is realized by the degree of methylation. As more attractant is sensed, the receptor complex is increasingly methylated. Likewise, if the bacterium senses less attractant, then the receptor complex is demethylated. The degree of methylation, therefore, provides a record of the past environment. Using this concept from control theory, receptor methylation is the bacterium's internal model of its environment. The bacterium decides whether to run or tumble by comparing the relative degree of receptor occupancy to the degree of receptor methylation: Are the current environmental conditions more favorable than they were? If so continue running, else tumble. This decision process demonstrates how bacteria are able to respond to gradients rather than the absolute state of their environment. Robustness is conferred by the mechanism of an integrator. The activity of the chemotaxis proteins (the parameters in the model) affects the dynamics of adaption, not the steady-state behavior. Although we expect that the dynamics have evolved for optimal chemotactic response, exact adaptation is a general network property and not a result of the specific parameterization. Robustness has also been demonstrated in networks such as phage $\lambda$ (80), circadian rhythms (81), and segment polarity development in *Drosophila* (82). These examples suggest that many regulatory properties are conferred solely by the network topology. We emphasize, however, that the network topology is not defined by the stoichiometry, or mass conservation, alone, as stoichiometry does not describe the regulatory properties of the network. Rather, the network is characterized by both the stoichiometry and the regulatory architecture.

Different species of bacteria employ different mechanisms for swimming such as *Rhodobacter sphaeriodes*, which has a single flagellum, *Sinorhizobium meliloti*, which is lophotrichously flagellated (83), and *Myxococcus xanthus*, which uses an unknown mechanism for gliding motility. One would expect, however, that *Bacillius subtilis*, which employs the same mechanism for swimming as *E. coli*, also possesses the same signal transduction pathway to mediate chemotaxis. Although the underlying design principles are the same and the regulatory motif is similar, the biochemical instantiation is different. *B. subtilis* possesses homologues to the *E. coli* proteins Tar (McpB), CheA, CheB, CheR, and CheY. There are also three additional proteins not found in *E. coli* that are implicated in *B. subtilis* chemotaxis: CheV (84), CheC, and CheD (85, 86). There is no analogue to the phosphatase CheZ in *B. subtilis*, though the receptor protein McpB contains a conserved region of CheZ (JR Kirby, personal communication). There are at least four functional differences between the chemotaxis pathway in *B. subtilis* and *E. coli* that suggest a different control strategy: (1) phosphorylated CheY (CheY-P) induces counterclockwise rotation rather than clockwise rotation in *B. subtilis* (87); (2) methylation is selective and the total number of methylated sites remains constant at steady state, whereas there is a net change in receptor methylation upon addition of attractants or repellents in *E. coli* (88); (3) CheY-P facilitates selective methylation (89); and (4) mutants lacking the receptor methylation in *B. subtilis* behave differently from the corresponding *E. coli* mutants (89). As the complete mechanism for chemotaxis in *B. subtilis* is currently unresolved, we can only speculate on the underlying mechanism. However, from the work of Yi and coworkers (77), we know adaptation requires an integrator. In *E. coli*, the integrator is realized biochemically by the degree of methylation. Whereas in *B. subtilis*, the integrator is realized apparently by the number of methyl groups at one site ($Gln^{371}$) relative to a second site ($Glu^{637}$).

If we understand the design principles, then we may postulate the mechanism with a limited number of experiments. In the case of chemotaxis, exact adaptation requires an integrator. Indeed, if this sort of zero-offset control is observed, then there is necessarily integral feedback. The chemotaxis pathway in *B. subtilis*, and in any other organism, must satisfy this constraint. When we study chemotaxis in *B. subtilis*, one of the first questions we need to answer is how does *B. subtilis* realize an integrator. We can also assume that both *B. subtilis* and *E. coli* employ similar regulatory motifs. As we have a model of *E. coli* chemotaxis, we can employ comparative analysis to facilitate our investigations and direct future experiments. This example illustrates the benefits of identifying necessary conditions in control and ways in which new control laws are realized.

## (Genetic) Switches: DNA Inversion and Cross-Regulatory Feedback Loops

In response to their environment, cells are capable of existing in many different physiological states. In many cases, different states are separated from each other

through the use of switches that allow cells to turn on and off different subsystems. Some of these switches, such as those underlying metabolic pathway choice, are reversible in that when the environmental stimulus is removed the cell returns to its basal state. Other switches, such as those underlying cellular differentiation, are irreversible or hysteretic such that the cell retains (at least for a while) the memory that the stimulus had been seen in the past. Switches are implemented in a large number of ways by cells; however, for most long-term changes switches involve the action of genetic networks. Even genetic switches are realized by a number of different control motifs. From an engineering perspective, switches are analogous to hybrid controllers utilizing propositional logic. A typical example is a failsafe mechanism that is triggered when either safety thresholds are violated or operating conditions change (26). These systems are ubiquitous in most control applications. Here we discuss two examples in order to show the different physical mechanisms by which genetic switches may be implemented. The first switch is the type-1 pili phase-variation switch in *E. coli*. This involves the mechanical rearrangement of the genome and is conceptually like an electrical knife-switch. The second switch is the λ phage lysis/lysogeny switch, which is implemented by cross-repressive feedback loops that look very much like an electronic latch. Both switches are stochastic, underlining the single-molecule nature of the DNA medium in which they are implemented.

The type-1 pili phase variation or *fim* switch in *E. coli* illustrates the dynamic nature of the genome. The *fim* switch in pathogenic *E. coli* controls the expression of type-1 pili used for adherence to and invasion of host tissue and is thought to be a virulence factor in urinary tract infections. Expression of *fim* is phase variable: Individual cells in the population alternate randomly between a piliated and nonpiliated state. This heterogeneity is necessary to guarantee a sufficiently large population of piliated cells for host tissue invasion, but not too large such that the colony cannot effectively spread or is vulnerable to the host's immune response to the pili antigen. The *fim* switch is temperature and nutrient sensitive, yielding apparently optimal heterogeneity for different environmental conditions (90). Pili expression is controlled by the orientation of a 314bp invertible element containing the promoter for the pili structural genes *fim A–fim H* (91). A proposed regulatory model for a minimal network, diagrammed in Figure 3, describes random phase variation in type 1 pili expression (DM Wolf & AP Arkin, unpublished results).

The phage λ switch controls the decision for lytic or lysogenic growth in response to extracellular signals (93). Figure 4 diagrams the core genetic switch, though the actual regulatory mechanism controlling the lysis/lysogeny decision is far more complicated (94). The switch is comprised of three operator sites controlling the promoters $P_{RM}$ and $P_R$. The Cro and CI dimers bind to the three operator sites and control the activity of the two promoters. The switch is realized by the two antagonistic feedback loops: The Cro dimer represses the expression of *cI* and the CI dimer represses the expression of the *cro*. The switch is also stochastic. Molecular fluctuations due to low protein concentrations leads to population heterogeneity even with identical initial conditions (cf. (94, 95)).

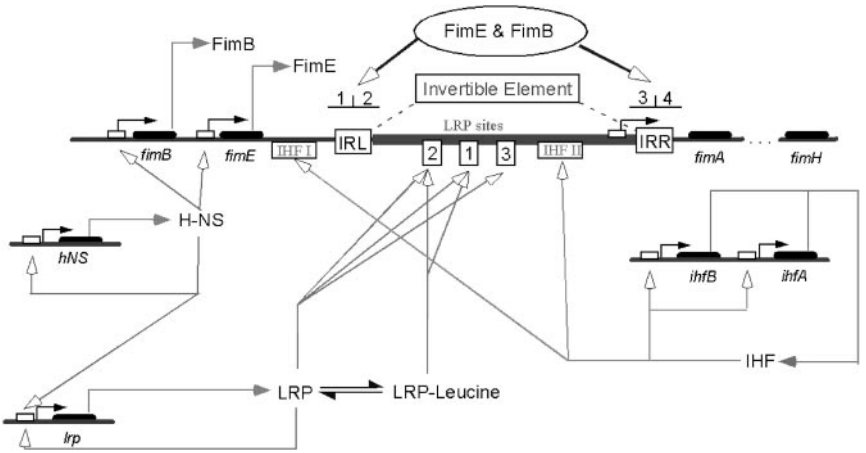



**Figure 3** The genetic network controlling the expression of type-1 pili in *E. coli*. The expression of the type-1 pili structure gene is controlled by the orientation of a 314bp invertible element. The switch is inverted independently by two recombinases FimB and FimE (92). Orientation is controlled by the expression of *fimE*, which preferentially inverts the switch to the off position, relative to the expression of *fimB*, which demonstrates no bias. The expression of *fimE* and *fimB* is controlled by the protein H-NS, whose expression is directed by a temperature sensitive repressor. The protein LRP, in dimer form, can bind to one of three sites on the invertible element facilitating inversion by bending the DNA segment. This effect is amplified when LRP binds with leucine. The protein IHF plays a supporting role in inversion. In this model, temperature and leucine concentration are the primary environmental controls through the expression of H-NS and the formation of the LRP-leucine complex. The nutritional content of the media also directs the expression of LRP.

One may compare the *fim* switch with the genetic switch controlling lysis and lysogeny in phage λ. Unlike the *fim* switch, the regulatory motif for the phage λ switch uses two divergent promoters regulated by two antagonist feedback loops. In particular, the *fim* switch is realized by physically inverting a DNA element, whereas the phage λ switch is realized by crossed feedback. Both the divergent promoters of phage λ and the DNA inversion of type-1 pili phase variation are common motifs. An example of another divergent promoter is the *pap* pili phase variation system. Another example of DNA inversion is flagellin phase variation in *Salmonella typhimurium*. Comparing and contrasting how control of these switches differ for different tasks should elucidate design tradeoffs and evolutionary selections of these different implementations.

Why have these two systems evolved different control motifs to regulate an environmentally sensitive developmental switch? One possible explanation is that the *fim* switch is either on or off, whereas the phage λ switch is fuzzy. Even

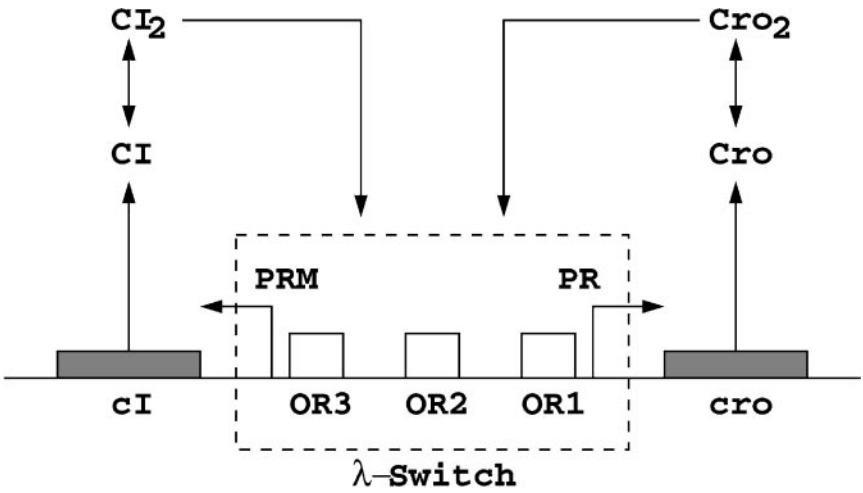



**Figure 4** The genetic network implementing the core bistability in the phage λ switch in *E. coli*. The three operator sites *OR1-3* control the promoters $P_{RM}$ and $P_R$. The Cro and CI dimers competitively bind to the three operator sites with different affinities. The Cro dimer represses the $P_{RM}$ promoter by binding to the operator site OR3 and, consequently, represses the synthesis of CI. The CI dimer activates the $P_{RM}$ promoter, represses the $P_R$ promoter by cooperatively binding to the operator sites OR1 and OR2, and consequently represses the synthesis of Cro. The net result is a pair of antagonistic feedback loops that yield a bistability. Control of this bistability is located nearby in the genome.

though there is a random element to the *fim* switch, the dynamics are relatively slow: Recombination occurs in at most 30% of the cells per generation. One might speculate that the *fim* motif is selected when concrete decisions are necessary, whereas the phage λ motif is selected for when the sole objective is to bias the fate. As many issues regarding the population dynamics of *E. coli* piliation are unresolved, the answer to this question is still open to debate.

## T7 Growth Cycle

The life cycle of a virus is analogous to an assembly line. Numerous components need to be assembled for the virus to replicate. Furthermore, the virus needs to schedule each task in order to maximize growth rate, as certain components are needed to assemble others. Scheduling problems arise also in most manufacturing processes. The question one seeks to answer is how to schedule each task in order to maximize efficiency.

Endy and coworkers (96) recently developed a detailed chemical-kinetic model of the infection life cycle of bacteriophage T7 in *E. coli* that predicts the intracellular dynamics of 57 phage gene products. T7 growth is partly regulated by the timing of insertion of phage DNA into the host cell. The life cycle may be

divided into three periods characterized by the expression of three different classes of gene. The class I genes initiate infection, the class II genes encode the phage DNA replication machinery, and the class III genes encode the phage particle and packaging proteins. The ordering of the phage genome, therefore, regulates phage replication. Furthermore, the transition to the second growth period is moderated by the T7 RNA polymerase. Once the T7 polymerase, encoded by a class I gene, is synthesized, it pulls the remainder of the genome into the host cell. Ordering of the genome provides a weak checkpoint in the T7 life cycle ensuring optimal growth. An analogous regulatory motif is used by *Salmonella typhimurium* in flagellar biosynthesis, where there is an ordered expression of three classes of genes (97). Unlike T7 though, the checkpoints are precise. Class II genes are not expressed until the genes in class I are functional. Likewise, Class III genes are not expressed until the genes in class II are functional.

In a subsequent paper, Endy and coworkers (98) investigated the effects of removing the scheduling enforced by the gene ordering on the genome by demonstrating computationally and experimentally how growth is affected by shuffling this order. Their model predicts that the natural ordering is nearly optimal, so long as the gene encoding the phage polymerase (gp1) is kept within the vicinity of the class I genes. Although their model occasionally predicted increased growth rates for certain permutations of gp1, the corresponding experiments using mutant phage appear to confirm the optimality of the natural ordering of the genome, though there are questions regarding the viability of these mutants. Remarkably though, a significant population ($\approx$50%) of the mutants yield progeny, though at a retarded growth rate, suggesting a degree of robustness in the organization of the genome. These results are consistent with many scheduling problems for manufacturing, where the solutions are not always unique and many suboptimal solutions exist.

Having used their models to investigate control of phage life by its endogenous network, Endy and coworkers used their kinetic model to investigate optimal antiviral strategies using antisense RNA targeted toward different T7 gene products. They were able to identify optimal targets for antisense strategies, thereby retarding the growth rate of T7. Interestingly, certain gene targets yielded increased growth as they exerted negative feedback on the phage RNA polymerase. In a subsequent paper, Endy & Yin (99) addressed the question of drug resistance. They demonstrated that certain gene targets, in particular the gene encoding the phage polymerase, cannot evolve resistance without first being at a selective disadvantage. In this sense, their drug target is robust. A novel feature of their analysis is that they considered the regulatory mechanisms in T7 infection. One limitation of their approach is that evolution does not proceed in a continuous fashion but rather in discrete steps, thus challenging the robustness of their design. Regardless, we believe their analysis is an important step toward the rational selection of targets for drug design. This series of papers demonstrates how understanding the control mechanisms of a particular biological system can also aid in design of external control strategies (pharmaceutical interventions).

## Cell Cycle

Another control issue that arises when scheduling many periodic processes are checkpoints: One needs to verify that each task is completed before the next task begins or the next batch is processed. These problems arise in the batch processing of specialty chemicals and pharmaceuticals and also in the design of asynchronous electronic circuits. These control systems are designed to prevent new tasks from beginning before the previous tasks are complete. The ordering of the genome in T7 enforces checkpoints between the three classes of genes. Another example of the use of checkpoints is in the cell cycle. This problem is periodic and is an example of how regulatory failure may lead to disease (100).

The eukaryotic cell cycle is divided into four phases: the $S$ phase (DNA replication), the $M$ phase (mitosis), and intervening phases $G_1$ and $G_2$. Being a free oscillator, precise, ordered control of the cell cycle is needed for viability, thus necessitating a series of checkpoints. The $G_1/S$ checkpoint ensures that the cell has divided since the previous round of DNA replication, that the cell is large enough to proceed with replication, and that the environmental conditions are suitable for mitosis. The $G_2/M$ checkpoint ensures that DNA replication is complete and any DNA damage has been repaired. Any failures in checkpoint control leads to cell cycle arrest. If the problem cannot be resolved, the cell undergoes apoptosis, or programmed cell death. Any breakdown in these failsafe mechanisms may lead to uncontrolled cell proliferation. Cancer, for example, is facilitated by a breakdown in cell cycle control. Understanding the mechanisms involved in regulating the cell cycle can facilitate the rational development of cancer therapies. We note that failsafe mechanisms analogous to cell cycle arrest and apoptosis are present in most industrial processes. These failsafe mechanisms ensure that any breakdown in control leads to a safe and ordered shutdown of the process.

Numerous regulatory models have been proposed for the cell cycle. Examples include the interaction of cyclin and cdc2 (101, 102), the $G_1/S$ transition (103, 104), the $G_2/M$ transition (105), and the cell cycle in fission (106, 107) and budding (108) yeast. A review of cell cycle modeling is given by Tyson (109). All of these models aim to understand the molecular mechanisms that control the cell cycle. Most models focus on the oscillatory dynamics in the cell cycle, though different mechanisms are used to explain the oscillations. The basic regulatory motif in the cell cycle involves the interaction of the cyclin dependent kinases (cdk's) and the associated cyclin proteins. Regulation is achieved through the formation of cdk-cyclin complexes, selective phosphorylation, and proteolytic degradation of the cyclins. Each mechanism is common to many regulatory schemes, though the particular organization appears unique to the control of the eukaryotic cell cycle. Although even in eukaryotes, this motif is more or less elaborate, Tyson is fond of pointing out that the yeast cell cycle is almost a subset of the human cell cycle: It is the "yeast within."

## Gene Expression: Autoregulation and Networks

Gene regulation does not illustrate a unique regulatory motif or define a specific control problem. Rather, it is ubiquitous in most facets of intracellular physiology. We do not aim to review the different mechanisms involved in regulating gene expression, but to illustrate how many aspects of this problem have been analyzed using tools from dynamical systems theory and also control theory. The work that we discuss is limited to autoregulation, where the gene expresses its own transcription factor, and simple gene networks, where typically one gene product regulates the expression of a second gene and vice versa. These problems demonstrate how regulation can be analyzed at many different levels of abstraction.

At least three approaches have been used to analyze the dynamics of genetic networks. One approach is to model the network using a Boolean logical network, where a gene is assumed to be either on or off (110). The network dynamics are realized by updating the logical state of each gene using a set of rules. For example, gene A is on at the current timestep only if gene B is off at the preceding time step. A second approach is to use differential equations to model gene activity and expression. For a comparison of the two approaches, the reader is directed to the survey article (20). The third approach is to model the gene network as a stochastic process (95). Differential equations are typically used to describe genetic networks as they are more detailed, though Boolean descriptions provide a simplified description that is often useful when first analyzing the network. Probabilistic descriptions are used to account for fluctuations in gene expression leading to population heterogeneity.

The dynamical properties of Boolean and differential equation models can often be analyzed directly using tools derived from dynamical systems theory. These studies characterize the dynamical properties of genetic networks. A typical result states that systems with negative feedback possess either a unique steady state or stable limit cycles, i.e. the system oscillates, and that a system with positive feedback may possess multiple steady states, i.e. the system acts as a genetic switch. Numerous articles have been written on the subject, and we highlight only a few related to differential equation models. These articles illustrate some of the results and techniques used to analyze genetic networks. Tyson & Othmer (111) considered the dynamics of gene regulation in enzyme synthesis with either end-product inhibition or induction. They analyzed, in a comprehensive manner, the dynamical properties of this network. Keller (112, 113) investigated the dynamical properties of a variety of control motifs using a genetic network with autoregulatory transcription factors. Wolf & Eeckman (114) investigated the dynamical properties of a genetic network using a simplified version of the model proposed by Shea & Ackers (115) for the phage $\lambda$ switch. Smolen and coworkers (116) investigated the dynamical properties of an autoregulatory genetic network and considered also the role of phosphorylation. In a separate paper (117), they considered the effects of a time delay between expression and binding of the transcription factor. Goldbeter (118) provides a comprehensive survey of the mechanisms underlying oscillations in genetic networks. Omholt and coworkers (119) analyzed the

dynamical properties of simple genetic networks and demonstrated how these networks can give rise to the genetic phenomena of additivity, dominance, and epistasis.

There are many different mechanisms for gene regulation. What factors have contributed to their selection? Comprehensive studies addressing this question have been done by Savageau and his colleagues in a series of articles spanning almost thirty years. One particular factor that they advocate is the parametric sensitivity of the network, one form of robustness (120). One may speculate that networks have evolved robustness (cf. (8)). In a subsequent series of articles (cf. (121–123)), Savageau and coworkers investigated different control motifs and speculated on their design principles. In one example (124), they considered kinetic proofreading in t-RNA aminoacylation and addressed the tradeoffs associated with accuracy versus cost in energy. In a separate series of articles, Savageau presented his demand theory for gene regulation (cf. (125–127)), which states that genes are regulated using an activator (e.g. *mal* operator) when function is in high demand while a repressor (e.g. *lac* operator) is used when function is in low demand.

In almost all cases of full analysis, the systems analyzed are small and consist of no more than two or three differential equations. Extending these results to large systems may be difficult. Often one needs to resort to numerical simulation to investigate the network behavior. This analysis may be time consuming as one needs to search a high dimensional parameter space. Occasionally, it is possible to reduce the complexity of the model (cf. (106)), but we suspect more often than not that direct analysis is limited to small systems. Regardless of this limitation, these problems present useful abstractions for analyzing gene expression in prokaryotes. One future challenge is to identify gene regulatory motifs in eukaryotes where the control mechanisms are far more complex (128, 129).

## FORWARD ENGINEERING: RATIONAL DESIGN OF GENETIC NETWORKS

A complementary approach to the methods described in the above section is the forward engineering approach. Rather than try to deduce the behavior of existing genetic networks, the next challenge is to engineer genetic networks in vivo to perform specific functions. The focus shifts then from deducing functional properties of naturally occurring genetic circuits toward engineering those properties de novo. Rationally engineering genetic circuits, therefore, provides perhaps the ultimate test of understanding (130). Engineered circuits also have potential application in medicine (e.g. gene therapy). Furthermore, engineered networks are interesting from a purely intellectual perspective.

In a series of papers described below, genetic networks have been engineered to perform certain functions (131–133). There is a long history of engineering genetic circuits, though the sophistication is increasing. In these examples, the genetic networks were engineered using a series of repressible promoters, with green fluorescent protein (GFP) used as a marker. We describe these networks in the

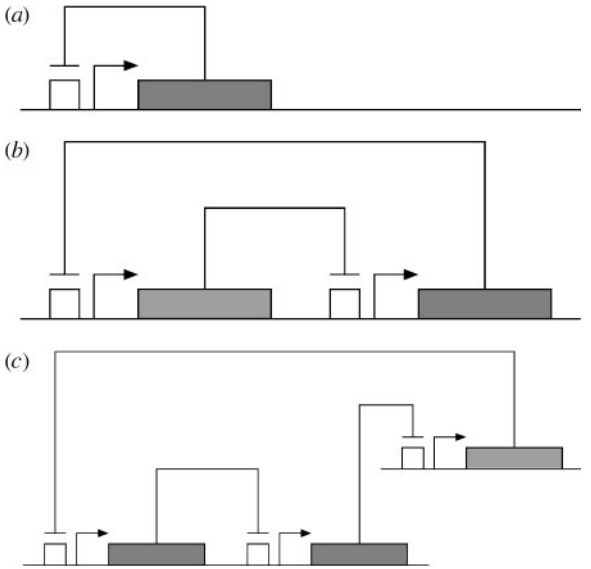



**Figure 5** (*a*) An example of an autoregulatory feedback loop, where a protein represses its own synthesis. (*b*) A genetic switch realized by two antagonistic feedback loops, where the first protein represses the synthesis of the second protein and the second protein represses the synthesis of the first protein. (*c*) An oscillator realized by three antagonistic feedback loops, where the first protein represses the synthesis of the second protein, the second protein represses the synthesis of the third protein, and the third protein represses the synthesis of the first protein.

order of their complexity. For information regarding the particulars of the designs, the reader is directed to the specific articles. Becskei & Serrano (131) constructed an autoregulatory network using a single repressible promoter (see Figure 5*a*). They demonstrated that negative feedback attenuates fluctuations in gene expression. Gardner and coworkers (132) constructed two different genetic toggle switches using a network consisting of two repressible promoters (see Figure 5*b*). Their design is similar to the phage λ switch, as it employs two antagonistic feedback loops. Both of their designs exhibit bistability. In one design, the switch is turned on by a pulse of isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) and turned off by a thermal pulse. In the second design, the switch is turned on by a pulse of IPTG and turned off by a pulse of anhydrotetracycline. Elowitz & Leibler (133) constructed an oscillator in vivo using a network consisting of three repressible promoters (see Figure 5*c*). Their design employs three antagonistic feedback loops. At least 40% of the transformed cells using their design were found to exhibit oscillatory behavior. The fact this number was not 100% indicates that the design was not perfect. In all three cases, simple mathematical models were used to facilitate the design. Deviations from model predictions can lead to new insights.

Forward engineering provides an alternate approach for elucidating the properties of biological networks by allowing us to directly tinker with their behavior. These simple networks, or control modules, provide complete knowledge of designed interactions and are fairly uncoupled from all other networks, though inevitably there is some coupling as these modules are integrated in the overall physiology of the cell. Most of the tools in systems and control theory are constructive rather than deductive, a natural consequence of being an engineering science. Using the forward engineering approach, one can postulate parsimonious biochemical and genetic circuits network using tools from systems and control engineering with robust properties and then design these networks in vivo. By designing and fabricating such networks and learning from our successes and failures, we can potentially identify design principles from scratch rather than trying to reverse-engineer them from nature. This sort of speculation is often impossible using computational or mathematical analysis alone, as there are always differences between predicted and observed results. We can further seek to design selective pressures for these engineered networks and attempt to ascertain their evolutionary viability. These selective pressures will also aid in determining why certain regulatory motifs are chosen instead of other motifs, even though their dynamic behavior appears identical.

## DISCUSSION

We have described some of the mechanisms used to regulate biological networks. Throughout our discussion, we have attempted to illustrate the analogies to engineered systems by describing different control motifs that commonly arise in engineering. We believe the analogies are clear, though how these systems are designed is not. As biologists, we are inevitably playing catch-up, but perhaps as engineers we can begin to close the gap. Many researchers have adopted this engineering perspective (cf. (134)) and have begun to enumerate design alternatives and constraints in biological systems. The theme of this article is control and the motifs used in regulation. We do not foresee a paradigm shift that will allow us to suddenly decipher the complexities of biological networks, but rather the gradual development of a set of design anecdotes or "rules of thumb." For example, it is extremely difficult to decipher the function of an integrated circuit simply by looking at the circuit diagram. Rather, one needs to understand the modular decomposition of the circuit and the principle used to design it. The engineer does not design an integrated system from scratch, but rather uses prefabricated functional components or unit operations. Inevitably, as the design matures, the system becomes more integrated and the functional boundaries of the components disappear. Still, artifacts of the modular design remain.

We will learn some of these principles by reverse-engineering the behavior of biological systems, whereas in other circumstances we will need to learn these principles by trial and error through designing our own networks. Already many

regulatory motifs have been identified. Furthermore, we are beginning to identify some of the components used in intracellular regulation. Enzymes can certainly act like switches or transistors either by allosteric changes in conformation or covalent modification (135). Certain pathways such as the MAPK cascade (136, 137) also act like switches. Other pathways such as two component phosphorelays (138) appear to act as filters or even amplifiers. By understanding the properties of these functional components and how they integrate into different regulatory motifs, we can begin to deduce network behavior as engineers. Many hurdles still exist and we conclude by discussing just a few.

## CHALLENGES AND OPPORTUNITIES

### Modeling

How do we model biological systems when we do not understand the complete mechanism? How do we conveniently abstract the problem? What level of detail is necessary to understand a given phenomenon? How do we incorporate multiple levels of resolution in our model? There is no unique answer to any of these questions. In these regards, modeling is still an art honed by many years of experience. Regardless, we need to answer these questions and distill the critical problems and issues. Only then can we advance the field of modeling in biology beyond a collection of anecdotes.

Another challenge in modeling biological systems is that even if we understand the mechanism, the parameters are rarely known. This problem plagues most modeling efforts with the exception of a few well characterized systems such as phage $\lambda$, the T7 growth cycle, and bacterial chemotaxis. Nevertheless, the modeling effort can answer important questions regarding the validity of the proposed mechanism. If we cannot find a plausible set of parameters that reproduces the behavior of the network, then the postulated mechanism is either incorrect or incomplete. This line of reasoning is employed in many cell cycle models. Another excellent example is a model of segment polarity development in *Drosophila* proposed by Dassow and coworkers (82). Their initial model was unable to produce asymmetric patterns, suggesting that their mechanism was incomplete. After revising their model by postulating new mechanisms, they were able to produce asymmetric patterns over a wide range of parameters, suggesting further that asymmetric patterning is a robust feature of network topology.

How do we directly couple modeling with experimentation? New measurement techniques are required to quantify intracellular protein concentrations and provide reliable estimates of gene expression. Statistical tools are needed to facilitate the modeling effort and direct experimentation. In particular, we require tools to infer regulatory mechanism from time-series data (cf. (139–141)) and also methods to design experiments for deducing the regulatory properties of interest. Given the vast array of biological databases, we require modeling environments

that provide convenient interfaces to access these databases. Some examples toward this goal include the Virtual Cell project (142) and the E-cell project (41).

Our knowledge of any system often comes in a wide variety of forms and at many different levels of granularity. How, for example, do we integrate macroscopic and microscopic information? Most simulation tools require that we operate at the highest level of resolution, a strategy that is neither optimal nor always feasible. Furthermore, what do we do when we understand the mechanisms for only part of the system and can, at best, represent the other parts using only phenomenological or qualitative descriptions? Given the complexity of biological systems and the need to operate at many different scales—from individual molecular events, such as transcription, to concentration dependent events, such as intercellular signalling, and macroscopic phenomena, such as motion—we need to be able to use hybrid, or mixed, models that will integrate knowledge at different scales and levels of detail. Many questions remain unanswered, from how to solve these systems to even how to formulate the model.

## Determinism versus Randomness

If the goal is only to understand the average behavior of a genetic network, then often one does need to account for stochastic fluctuations. If we seek, however, to describe population heterogeneity, then we need to account for stochastic fluctuations associated with random noise due to low molecular concentrations. As we have already mentioned, random noise accounts for populations heterogeneity in both *E. coli* piliation and phage $\lambda$ lysis/lysogeny. One can attribute these variations to fluctuations in gene expression (95, 143). Likewise, random variations are also implicated in a wide variety of swimming behaviors observed in bacteria (144, 145).

Although probabilistic descriptions provide accurate representations of the dynamics at low concentrations, they are far more difficult to analyze than deterministic systems. Stochastic systems do not possess steady states but rather stationary distributions describing the likelihood of a given concentration. Enumerating the stationary distributions and characterizing their stability is far more difficult than in deterministic systems, as the differential equations are replaced by the chemical master equation (146, 147). Even calculating a stationary distribution is difficult. Often one is only able to calculate a single realization of a stochastic system, akin to drawing a single card rather than determining the odds of drawing an ace. One can still calculate the distribution using Monte Carlo strategies, but these do not allow us to determine the stability properties of the distributions or systematically identify possible bifurcations. In deterministic systems there is no difference between a realization and a distribution as the same series of events happen for a given initial condition.

Given the need for probabilistic descriptions and also the difficulty associated with analyzing stochastic systems, is it possible to identify when a probabilistic description is necessary? We desire simulation tools that will adapt to our

demands. As probabilistic descriptions are not always necessary to answer the questions being asked, an adaptive framework or algorithm that can recognize the need for a probabilistic description, incorporate both deterministic and stochastic descriptions, and switch between the two would substantially improve our ability to simulate and analyze biological systems.

## Biophysical Constraints

Most models of biological networks assume that the intracellular proteins are homogeneously distributed. The actual intracellular environment, however, is heterogeneously distributed, and effects due to compartmentalization (148), molecular crowding (149, 150), enzyme complexes (151–154), and diffusion (117, 155) may play a significant role in intracellular regulation. Consequently, we cannot limit our analysis to the biochemistry; we need to take into account and understand the role of biophysical constraints. One example is receptor clustering (156). In bacterial chemotaxis, the receptors are polarly localized in the cell (157). What is remarkable about receptor clustering is that it was first hypothesized that the receptors should be uniformly distributed on surface of the cell (63). Receptor clustering has since been hypothesized to play a role in the sensitivity of the chemotaxis pathway (158–160).

A challenge in considering biophysical constraints is that the analysis is far more complicated. In particular, we need to account for spatio-temporal behavior, such as calcium waves, and complicated geometries. Sometimes it is possible to approximate the problem using compartment models or time delays. However, compartment models ignore surface geometry such as membrane folds leading to a nonuniform distribution of the surface proteins. One caveat with time delays is that they are difficult to solve numerically and are known to introduce instability in feedback loops. One needs to ensure that these effects are not due to numerical instabilities nor are artifacts of a fixed time delay.

## The Path Forward

As we begin to investigate increasingly more complex systems, it is unlikely that we will be able to continue using an undirected, reductionist approach. To date many regulatory models have been obtained by breaking the system into its elementary pieces and then trying to reassemble the pieces back together. We are now being confronted with so many pieces that we will unlikely be able to reassemble the system without multiple levels of abstraction. An alternate approach is to attack the problem from the perspective of evolution and tinker with known modules or regulatory motifs (161). Our aim in this review is to point out some of the design principles and motifs used to regulate genetic and biochemical networks. As our knowledge expands, we hope to assemble a catalog, or rather a toolbox, of functional modules and regulatory motifs for the modeler to tinker with. Ultimately, what separates biology from the physical sciences is evolution. We cannot forget that biological systems are the product of an evolutionary process; thus comparing convergently and divergently evolved motifs should provide information on the selection for and fitness criterion of cellular function.

## ACKNOWLEDGMENTS

This work was supported by grants from DARPA, ONR, and DOE. The authors wish to thank John Kirby for many helpful discussions on chemotaxis in *B. subtilis* and Denise Wolf for many helpful discussions on the *fim* switch in *E. coli*. We also thank Janet Jacobsen and Keith Erickson for their comments and suggestions.



## LITERATURE CITED

1. Carlson JM, Doyle J. 1999. Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys. Rev. E* 60:1412–27
2. Weinberg RA. 1996. How cancer arises. *Sci. Am.* 275:62–70
3. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular biology to modular cell biology. *Nature* 402:C47–C51
4. Lauffenburger DA. 2000. Cell signaling pathways as control modules: complexity for simplicity. *Proc. Natl. Acad. Sci. USA* 97:5031–33
5. Asthagiri AR, Lauffenburger DA. 2000. Bioengineering models of cell signalling. *Annu. Rev. Biomed. Eng.* 2:31–53
6. Bhalla US, Iyengar R. 1999. Emergent properties of networks of biological signaling pathways. *Science* 283:381–87
7. Bhalla US, Iyengar R. 1999. Complexity in biological signaling systems. *Science* 284:92–96
8. Kirschner M, Gerhart J. 1998. Evolvability. *Proc. Natl. Acad. Sci. USA* 95:8420–27
9. Wiener N. 1961. *Cybernetics or Control and communication in the animal and machine*. Cambridge: MIT Press
10. Kalman RE. 1968. New developments in systems theory relevant to biology. In *Systems Theory and Biology* (MD Mesarovi'c, ed.), (New York), 222–32, Springer-Verlang
11. DeResi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680–86
12. Brown PO, Botstein D. 1999. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* 21(Suppl):33–37
13. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, et al. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95:717–28
14. Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. 1999. Systematic determination of genetic network architecture. *Nat. Genet.* 22:281–85
15. Karp PD, Riley M, Paley SM, Pellegrini-Toole A, Krummenacker M. 1999. Eco Cyc: Encyclopedia of *Escherichia coli* genes and metabolism. *Nucleic Acids Res.* 27:55–58
16. Ogata H. 1999. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27:29–34
17. D'Souza M, Romine MF, Maltsev N. 2000. SENTRA, a database of signal transduction proteins. *Nucleic Acids Res.* 28:335–36
18. Overbeek R, Larsen N, Pusch GD, Dsouza M, Selkov E, et al. 2000. WIT: integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Res.* 28: 123–25
19. Palsson B. 2000. The challenges of *in silico* biology. *Nat. Biotech.* 18:1147–50

20. 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
21. Smolen P, Baxter DA, Byrne JH. 2000. Mathematical modeling of gene networks. *Neuron* 26:567–80
22. McAdams HH, Arkin A. 1998. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* 27:199–224
23. Gombert AK, Nielsen J. 2000. Mathematical modelling of metabolism. *Curr. Opin. Biotechnol.* 11:180–86
24. Arkin AP. 2000. Signal processing biochemical reactions networks. In *Self-organized Biological Dynamics and Nonlinear Control*, ed. J Walleczek, pp. 112–144. UK: Cambridge University Press
25. Luyben WL, Tyreus BD, Luyben ML. 1998. *Plantwide Process Control*. New York: McGraw-Hill
26. Ogunnaike BA, Ray WH. 1994. Process Dynamics, Modeling, and Control. New York: Oxford University Press
27. Bailey JE. 1991. Towards a science of metabolic engineering. *Science* 252: 1669–75
28. Koffas M, Roberge C, Lee K, Stephanopoulos G. 1999. Metabolic engineering. *Annu. Rev. Biomed. Eng.* 1:535–57
29. Stephanopolous G, Vallino JJ. 1991. Network rigidity and metabolic engineering in metabolite overproduction. *Science* 252:1675–81
30. Westerhoff HV. 1995. Subtlety in control—metabolic pathway engineering. *Trends Biotechnol.* 13:242–44
31. Fell DA. 2000. Signal transduction and the control of expression of enzyme activity. *Advan. Enzyme Regul.* 40:35–46
32. Ouzounis CA, Karp PD. 2000. Global properties of the metabolic map of *Escherichia coli*. *Genome Res.* 2000:568–678
33. Bailey JE. 1998. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnol. Prog.* 14:8–20
34. Varner J, Ramkrishna D. 1999. Mathematical models of metabolic pathways. *Curr. Opin. Biotechnol.* 10:146–50
35. Giersch C. 2000. Mathematical modelling of metabolism. *Curr. Opin. Plant Biol.* 3:249–53
36. Arkin A, Ross J. 1994. Computational functions in biochemical reaction networks. *Biophys. J.* 67:560–78
37. Joshi A, Palsson BO. 1989. Metabolic dynamics in the human red cell. Part I–a comprehensive kinetic model. *J. Theoret. Biol.* 141:515–28
38. Palsson BO, Lee I-D. 1993. Model complexity has a significant effect on the numerical value and interpretation of metabolic sensitivity coefficients. *J. Theoret. Biol.* 161:299–315
39. Ni TC, Savageau MA. 1996. Application of biochemical systems theory to metabolism in human red blood cells. *J. Biol. Chem.* 271:7927–41
40. Ni TC, Savageau MA. 1996. Model assessment and refinement using strategies from biochemical systems theory: application to metabolism in human red blood cells. *J. Theoret. Biol.* 179:329–68
41. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, et al. 1999. E-CELL: software environment for whole-cell simulation. *Bioinformatics* 15:72–84
42. Varma A, Palsson BO. 1993. Metabolic capabilities of *Escherichia coli*: II. Optimal growth patterns. *J. Theoret. Biol.* 165:503–22
43. Edwards JS, Palsson BO. 1999. Systems properties of the *Haemophilus influenzae* rb metabolic genotype. *J. Biol. Chem.* 97:17410–16
44. Edwards JS, Palsson BO. 2000. The *Escherichia coli* mg1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* 97:5528–33
45. Ahuja RK, Magnanti TL, Orlin JB. 1993. *Network Flows: Theory, Algorithms, and*

*Applications*. Englewood, NJ: Prentice-Hall
46. Clarke BL. 1980. Stability of complex reaction networks. *Adv. Chem. Phys.* 43:1–215
47. Schilling CH, Palsson BO. 1998. The underlying pathway structure of biochemical reaction networks. *Proc. Natl. Acad. Sci. USA* 95:4193–98
48. Schilling CH, Letscher D, Palsson BO. 2000. Theory of the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theoret. Biol.* 203:229–48
49. Schuster S, Fell DA, Dandekar T. 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotech.* 18:326–32
50. Horn F, Jackson R. 1972. General mass action kinetics. *Arch. Ration. Mech. Anal.* 47:81–116
51. Feinberg M. 1995. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Ration. Mech. Anal.* 47:311–70
52. Kacser H, Burns JA. 1973. The control of flux. *Symp. Soc. Exp. Biol.* 27:65–104
53. Heinrich R, Rapoport TA. 1974. A linear steady-state treatment of enzymatic chains: general properties and effector strength. *Eur. J. Biochem.* 42:89–95
54. Reder C. 1988. Metabolic control theory: a structural approach. *J. Theoret. Biol.* 125:175–201
55. Fell DA. 1992. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem. J.* 286:313–30
56. Heinrich R, Schuster S. 1996. *The Regulation of Cellular Systems*. New York: Chapman & Hall
57. Fell D. 1997. *Understanding the Control of Metabolism*. London: Portland
58. Szyperski T, Glaser RW, Hochuli M, Fiaux J, Sauer U, et al. 1999. Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fraction $^{13}C$ labelling and two-dimensional NMR spectroscopy. *Metabol. Eng.* 1:189–97
59. Schmidt K, Norregaard LC, Pedersen B, Meissner A, Duus JO, et al. 1999. Quantification of intracellular metabolic flux from fractional enrichment and $^{13}C-^{13}C$ coupling constraints on the isotopomer distribution in labeled biomass components. *Metabol. Eng.* 1:166–79
60. Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, et al. 1999. Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *J. Bacteriol.* 181:6679–88
61. Roscher A, Kruger NJ, Ratcliffe RG. 2000. Strategies for metabolic flux analysis in plants using isotope labelling. *J. Biotechnol.* 77:81–102
62. Berg HC. 1975. Chemotaxis in bacteria. *Annu. Rev. Biophys. Bioeng.* 4:119–36
63. Berg HC, Purcell EC. 1977. Physics of chemoreception. *Biophys. J.* 20:193–215
64. Koshland DE. 1977. A response regulator model in a simple sensory system. *Science* 196:1055–63
65. Koshland DE, Goldbeter A, Stock JB. 1982. Amplification and adaptation in regulatory and sensory systems. *Science* 217:220–25
66. Parkinson JS. 1993. Signal transduction schemes of bacteria. *Cell* 73:857–71
67. Stock JB. 1994. Adaptive responses in bacterial chemotaxis. In *Regulation of Cellular Signal Transduction by Desensitization and Amplification*, ed. DR Sibley, MD Houslay, pp. 3–24. New York: Wiley
68. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA. 1997. The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell. Dev. Biol.* 13:457–512
69. Goldbeter A, Koshland DE. 1982. Simple molecular model for sensing and adaptation based on receptor modification with

application to bacterial chemotaxis. *J. Mol. Biol.* 161:395–416
70. Asakura S, Honda H. 1984. Two-state model for bacterial chemoreceptor proteins: the role of multiple methylation. *J. Mol. Biol.* 176:349–67
71. Bray D, Bourret RB, Simon MI. 1993. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol. Biol. Cell* 4:469–82
72. Hauri DC, Ross J. 1995. A model of excitation and adaptation in bacteria chemotaxis. *Biophys. J.* 68:708–22
73. Spiro PA, Parkinson JS, Othmer HG. 1997. A model of excitation and adaptation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA* 94:7263–68
74. Barkai N, Leibler S. 1997. Robustness in simple biochemical networks. *Nature* 387:913–17
75. Morton-Firth CJ, Shimizu TS, Bray D. 1999. A free-energy-based stochastic simulation of the Tar receptor complex. *J. Mol. Biol.* 286:1059–74
76. Alon U, Surette MG, Barkai N, Leibler S. 1999. Robustness of bacterial chemotaxis. *Nature* 397:168–71
77. Yi TM, Huang Y, Simon MI, Doyle J. 2000. Robust perfect adaptation in bacterial chemotaxis through integral feedback. *Proc. Natl. Acad. Sci. USA* 99:4649–53
78. Francis BA, Wonham WM. 1976. The internal model principle of control theory. *Automatica* 12:457–65
79. Wonham WM. 1985. *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag, 3rd ed.
80. Little JW, Shepley DP, Wert DW. 1999. Robustness of a gene regulatory circuit. *EMBO J.* 18:4299–307
81. Barkai N, Leibler S. 2000. Biological rhythms: circadian clocks limited by noise. *Nature* 403:267–68
82. Dassow Gvon, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust development module. *Nature* 406:188–92
83. Armitage JP, Schmitt R. 1997. Bacterial chemotaxis: *Rhodobacter sphaeroides* and *Sinorhizobium meliloti*: variations on a theme? *Microbiology* 143:3671–82
84. Fredrick KL, Helmann JD. 1994. Dual chemotaxis signalling pathways in *Bacillus subtilis*: a sigma D-dependent gene encodes a novel protein with both CheW and CheY homologous domain. *J. Bacteriol.* 176:2727–35
85. Rosario MML, Kirby JR, Bochar DA, Ordal GW. 1995. Chemotactic methylation and behavior in *Bacillus subtilis*: role of two unique proteins, CheC and CheD. *Biochemistry* 34:3823–31
86. Rosario MML, Ordal GW. 1996. CheC and CheD interact to regulate methylation of *Bacillus subtilis* methyl-accepting chemotaxis proteins. *Mol. Microbiol.* 21:511–18
87. Bischoff DS, Bourret RB, Kirsch ML, Ordal GW. 1993. Purification and characterization of *Bacillus subtilis* CheY. *Biochemistry* 32:9256–61
88. Zimmer MA, Tiu J, Collins MA, Ordal GW. 2000. Selective methylation changes on the *Bacillus subtilis* chemotaxis receptor McpB promote adaptation. *J. Biol. Chem.* 275:24264–72
89. Kirby JR, Saulmon MM, Kristich CJ, Ordal GW. 1999. CheY-dependent methylation of the asparagine receptor, McpB, during chemotaxis in *Bacillus subtilis*. *J. Biol. Chem.* 274:11092–100
90. Gally DL, Brogan JA, Eisenstein BI, Blomfield IC. 1993. Environmental regulation of the *fim* switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J. Bacteriol.* 175:6186–93
91. Abraham JM, Freitag CS, Clements JR, Eisenstein BI. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 82:5724–27
92. McClain MS, Blomfield IC, Eisenstein BI. 1991. Roles of fimB and fimE in site-specific DNA inversion associated with phase variation of type 1 fimbriae in

*Escherichia coli*. *J. Bacteriol.* 173:5308–14
93. Ptashne M. 1992. *A Genetic Switch: Phage λ and Higher Organisms*. Cambridge, Mass.: Cell Press: Blackwell Scientific, 2nd ed.
94. Arkin A, Ross J, McAdams HH. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* 149:1633–48
95. McAdams HH, Arkin A. 1997. Stochastic mechanisms for gene expression. *Proc. Natl. Acad. Sci. USA* 94:814–19
96. 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
97. Hughes KT, Mathee K. 1998. The anti-sigma factors. *Annu. Rev. Microbiol.* 52: 231–86
98. Endy D, You L, 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. Endy D, Yin J. 2000. Toward antiviral strategies that resist viral escape. *Antimicrob. Agents Chemother.* 44:1097–99
100. Hartwell LH, Kastan MB. 1994. Cell cycle control and cancer. *Science* 266:1821–28
101. Tyson JJ. 1991. Modeling the cell division cycle: cdc2 and cyclin interactions. *Proc. Natl. Acad. Sci. USA* 88:7328–32
102. Goldbeter A. 1991. A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. *Proc. Natl. Acad. Sci. USA* 88:9107–11
103. Obeyesekere MN, Herbert JR, Zimmerman SO. 1995. A model of *G*1 phase of the cell cycle incorporating cyclin E/cdk2 complex and retinoblastoma protein. *Oncogene* 11:1199–205
104. Hatzimanikatis V, Lee KH, Bailey JE. 1999. A mathematical description of regulation of the G1-S transition of the mammalian cell cycle. *Biotechnol. Bioeng.* 65:631–37
105. Aguda BD. 1999. Instabilities in phosphorylation-dephosphorylation cascades and cell cycle checkpoints. *Oncogene* 18: 2846–51
106. Novak B, Tyson JJ. 1997. Modeling the control of DNA replication in fission yeast. *Proc. Natl. Acad. Sci. USA* 94:9147–52
107. Sveiczer A, Csikasz-Nagy A, Gyorffy B, Tyson JJ, Novak B. 2000. Modeling the fission yeast cell cycle: quantized cycle times in wee1$^-$ cdc25Δ mutant cells. *Proc. Natl. Acad. Sci. USA* 97:7865–70
108. Chen KC, Csikaz-Nagy A, Gyorffy Y, Val J, Novak B, Tyson JJ. 2000. Kinetic analysis of a molecular model of the budding yeast cell cycle. *Mol. Biol. Cell* 11:369–91
109. Tyson JJ. 1999. Models of cell cycle control in eukaryotes. *J. Biotechnol.* 71:239–44
110. Thieffry D, Thomas R. 1998. Qualitative analysis of gene networks. *Pac. Symp. Biocomput.*, Maui, HI, pp. 77–88
111. Tyson JJ, Othmer HG. 1978. The dynamics of feedback control circuits in biochemical pathways. *Prog. Theor. Biol.* 5:1–62
112. Keller AD. 1994. Specifying epigenetic states with autoregulatory transcription factors. *J. Theoret. Biol.* 170:175–81
113. Keller AD. 1995. Model genetic circuits encoding autoregulatory transcription factors. *J. Theoret. Biol.* 172:169–85
114. Wolf DM, Eeckman FH. 1998. On the relationship between genomic regulatory element organization and gene regulatory dynamics. *J. Theoret. Biol.* 195:167–86
115. Shea MA, Ackers GK. 1985. The $O_R$ control system of bacteriophage lambda: a physical-chemical model of gene regulation. *J. Mol. Biol.* 181:211–30
116. Smolen P, Baxter DA, Byrne JH. 1998. Frequency selectivity, multistability, and oscillations emerge from models of

genetic regulatory systems. *Am. J. Physiol.* 277:C531–C542
117. Smolen P, Baxter DA, Byrne JH. 1999. Effects of macromolecular transport and stochastic fluctuations on dynamics of genetics regulatory systems. *Am. J. Physiol.* 277:C777–C790
118. Goldbeter A. 1996. *Biochemical Oscillations and Cellular Rhythms: the Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge, NY: Cambridge Univ. Press
119. Omholt SW, Plahte E, Oyehaug L, Xiang K. 2000. Gene regulatory networks generating the phenomena of additivity, dominance, and epistasis. *Genetics* 155: 969–80
120. Savageau MA. 1971. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* 229:542–44
121. Savageau MA, Jacknow G. 1979. Feedforward inhibition in biosynthetic pathways: inhibition of the aminoacyl-tRNA synthetase by intermediates of the pathway. *J. Theoret. Biol.* 77:405–25
122. Savageau MA. 1985. A theory of alternative designs for biochemical control systems. *Biomed. Biochim. Acta* 44:875–80
123. Hlavacek WS, Savageau MA. 1997. Completely uncoupled and perfectly coupled gene expression in repressible systems. *J. Mol. Biol.* 256:538–58
124. Savageau MA, Freter RR. 1979. On the evolution of accuracy and cost of proofreading t-RNA aminoacylation. *Proc. Natl. Acad. Sci. USA* 76:4507–10
125. Savageau MA. 1974. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* 252:546–49
126. Savageau MA. 1977. Design of molecular control mechanisms and the demand for gene expression. *Proc. Natl. Acad. Sci. USA* 74:5647–51
127. Savageau MA. 1998. Demand theory of gene regulation. I. Quantitative development of theory. *Genetics* 149:1665–76
128. Yuh C-H, Bolouri H, Davidson EH. 1998. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* 279:1896–902
129. Arnone MI, Davidson EH. 1997. The hardwiring of development: organization and function of genomic regulatory systems. *Development* 124:1851–64
130. McAdams HH, Arkin A. 2000. Towards a circuit engineering discipline. *Curr. Biol.* 10:R318–R320
131. Becskei A, Serrano L. 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590–93
132. Gardner TS, Cantor CR, Collins JJ. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–42
133. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–38
134. McAdams HH, Shapiro L. 1995. Circuit simulation of genetic networks. *Science* 269:650–56
135. Goldbeter A, Koshland DE. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA* 78:6840–44
136. Huang C-YF, Ferrell JE. 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci. USA* 93:10078–83
137. Ferrell JE, Machleder EM. 1998. The biochemical basis of an all-or-none cell fate switch in *Xenupos ooctyes*. *Science* 280:895–98
138. Stock AM, Robinson VL, Goudreau PN. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* 69:183–215
139. Arkin A, Ross J. 1995. Statistical construction of chemical reaction mechanisms from measured time-series. *J. Phys. Chem.* 99:970–79
140. Arkin A, Shen P, Ross J. 1997. A test case of correlation metric construction of a reaction pathway from measurements. *Science* 277:1275–79

141. You L, Yin J. 2000. Patterns of regulation from mRNA and protein time series. *Metabol. Eng.* 2:210–17
142. Schaff J, Fink CC, Slepchenko B, Carson JH, Loew LM. 1997. A general computational framework for modeling cellular structure and function. *Biophys. J.* 73:1135–45
143. Ko MSH. 1992. Induction mechanism of a single gene molecule: stochastic or deterministic. *BioEssays* 14:341–46
144. Spudich JL, Koshland DE. 1976. Nongenetic individuality: chance in the single cell. *Nature* 262:467–71
145. Levin MD, Morton-Firth CJ, Abouhamad WN, Bourret RB, Bray D. 1998. Origins of individual swimming behavior in bacteria. *Biophys. J.* 74:175–81
146. Gillespie DT. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81:2340–61
147. Gillespie DT. 1992. A rigorous derivation of the chemical master equation. *Physica A* 188:404–25
148. Bakker BM, Mensonides FI, Teusink B, van Hoek P, Michels PA, Westerhoff HV. 2000. Compartmentation protects trypanosomes from the dangerous design of glycolysis. *Proc. Natl. Acad. Sci. USA* 97:2087–92
149. Zimmerman SB, Milton AP. 1993. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* 22:27–65
150. Rohwer JM, Postma PW, Kholodenko BN, Westerhoff HV. 1998. Implications of macromolecular crowding for signal transduction and metabolite channeling. *Proc. Natl. Acad. Sci. USA* 95:10547–52
151. Bray D, Lay S. 1997. Computer-based analysis of the binding steps in protein complex formation. *Proc. Natl. Acad. Sci. USA* 94:13493–98
152. Levchenko A, Bruck J, Sternberg PW. 2000. Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proc. Natl. Acad. Sci. USA* 97:5818–23
153. Burack WR, Shaw AS. 2000. Signal transduction: hanging on a scaffold. *Curr. Opin. Cell Biol.* 12:211–16
154. Ferrell JE. 2000. What do scaffold proteins really do? *Science's STKE* Perspectives (3 Oct 2000)
155. Ceresa BP, Schmid SL. 2000. Regulation of signal transduction by endocytosis. *Curr. Opin. Cell Biol.* 12:204–10
156. Bray D. 1998. Signalling complexes: biophysical constraints on intracellular communication. *Annu. Rev. Biophys. Biomol. Struct.* 27:59–75
157. Maddock JR, Shapiro L. 1993. Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science* 259:1717–23
158. Bray D, Levin M, Morthon-Firth CJ. 1998. Receptor clustering as a cellular mechanism to control sensitivity. *Nature* 393:85–88
159. Stock J, Da Re S. 1999. A receptor scaffold mediates stimulus-response coupling in bacterial chemotaxis. *Cell Calcium* 26:157–64
160. Shimizu ST, Le Novere N, Levin MD, Beavil AJ, Sutton BJ, Bray D. 2000. Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. *Nat. Cell Biol.* 2:792–96
161. Jacob F. 1977. Evolution and tinkering. *Science* 196:1161–66