# Chapter 23

# Dynamics and Control of Biological Systems

Previous chapters have emphasized the design of controllers for chemical process systems, as well as for biomedical systems (Chapter 22). In this chapter, we consider the analysis of intrinsically closed-loop systems that exist in biological circuits, from gene level through cellular level. There is no external controller to be synthesized; rather, the tools that were developed in the first half of this textbook are applied to the analysis of networks that exploit principles of feedback and feedforward control. These biophysical networks display the same rich character as those encountered in process systems engineering: multivariable interactions, complex dynamics, and nonlinear behavior. Examples are drawn from gene regulatory networks, as well as from protein signal transduction networks, with an emphasis on the role of feedback. A glossary of key technical terms is provided at the end of the chapter.

## 23.1 SYSTEMS BIOLOGY

Biophysical networks are remarkably diverse, cover a wide spectrum of scales, and are characterized by a range of complex behaviors. These networks have attracted a great deal of attention at the level of gene regulation, where dozens of input connections may characterize the regulatory domain of a single gene in a eukaryote, as well as at the protein level, where hundreds to thousands of interactions have been mapped in protein interactome diagrams that illustrate the potential coupling of pairs of proteins (Campbell and Heyer, 2007; Barabasi, 2004). However, these networks also exist at higher levels, including the coupling of individual cells via signaling molecules, the coupling of organs via endocrine signaling, and, ultimately, the coupling of organisms in ecosystems. The biochemical notion of signaling is discussed in Section 23.3. To elucidate the mechanisms employed by these networks, biological experimentation and intuition by themselves are insufficient. Instead, investigators characterize dynamics via mathematical models and apply control principles, with the goal of guiding further experimentation to better understand the biological network (Kitano, 2002). Increased understanding can facilitate drug discovery and therapeutic treatments.

A simple example that illustrates the roles of feedback and feedforward control in nature is the heat shock response exhibited by simple bacteria (El-Samad et al., 2006), as illustrated in Fig. 23.1. When the organism experiences an increase in temperature, it leads to the misfolding of protein, which disrupts a number of metabolic processes. One of the immediate effects of a heat disturbance is the feedforward activation of a component, $\sigma^{32}$, which turns on the transcription process for a pair of genes (FtsH and DnaK) that facilitates the repair mechanism for a misfolded protein. In particular, the FtsH gene is a protease, which catalyzes the destruction of the improperly folded protein. In parallel, and independently, the protein product of one of those genes (DNAK) monitors the state of protein misfolding. It binds to $\sigma^{32}$ and releases $\sigma^{32}$ when misfolded protein is detected, leading to feedback activation of DnaK transcription.

A second example of networked biological control is the circadian clock, which coordinates daily physiological behaviors of most organisms. The word *circadian* comes from the Latin for "approximately one day," and the circadian clock is vital to regulation of metabolic processes in everything from simple fungi to humans. The mammalian circadian master clock resides in the hypothalamus region of the brain (Reppert and Weaver, 2002). It is a network of multiple autonomous noisy oscillators, which communicate via neuropeptides to synchronize and form a coherent oscillator (Herzog et al., 2004; Liu et al., 2007). At the core of the clock is a gene regulatory network, in which approximately six classes of genes are regulated through an elegant array of

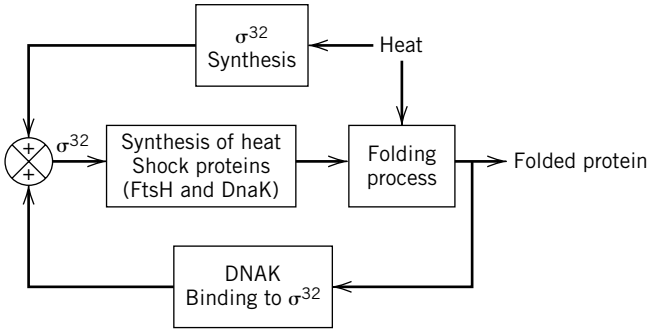


**Figure 23.1** Feedback and feedforward control loops that regulate heat shock in bacteria (modified from El-Samad, et al., 2006) (positive feedback is common in biological systems).

time-delayed negative feedback circuits (see Figure 23.2, which illustrates two of those six gene classes). The activity states of the proteins in this network are modulated (activated/inactivated) through a series of chemical reactions, including phosphorylation and dimerization. These networks exist at the subcellular level. Above this layer is the signaling that leads to a synchronized response from the population of thousands of clock neurons in the brain. Ultimately, this coherent oscillator then coordinates the timing of daily behaviors, such as the sleep/wake cycle. An interesting property of the clock is that, under conditions of constant darkness, the clock free-runs, with a period of approximately 24 h (i.e., "circa"), such that its internal time, or phase, drifts away from that of its environment. However, in the presence of an entraining cue (i.e., forcing signal, such as the rising and setting of the sun), the clock locks on to the period of that cue (Boulos et al., 2002; Dunlap et al., 2004; Daan and Pittendrigh, 1976). This gives rise to a precise 24-h period for the oscillations in protein concentrations for the feedback circuit in Fig. 23.2.

The Central Dogma tenet that most students learn in high school biology is a good starting point to understand these complex networks. Information in the cell is encoded in the DNA, and that information is expressed by the gene to produce messenger RNA. The mRNA is translated into a protein, which is one of the key building blocks of cells and which plays a critical role in cellular regulation. This form of the Central Dogma suggests a serial process, or a feedforward process, in which the genetic code influences the outcome (protein level and protein function). In some of the early publicity surrounding the Human Genome project, this type of logic was pervasive, and there was an understanding in some circles that the "parts list" (genetic code) would illuminate the cause of diseases. An engineer immediately recognizes the flaw in this logic: by analogy, if one were provided with the raw materials list for an aircraft (sheet metal, nuts, bolts, rivets, etc.), it would be an impossible leap to conclude anything about the principles of aerodynamics. Critical missing elements are the manner in which the parts are arranged into a network and, more important, how the components are controlled (or regulated). The same reasoning applies equally to biological networks as well, and this notion of the systems perspective has driven current research in systems biology.

According to the Central Dogma tenet, the additional layers of control and regulation that are mentioned in the preceding paragraph can be incorporated schematically, as shown in Fig. 23.3. Feedback control plays a key role in (i) regulation of the transcription event; (ii) processing of the RNA, including its stability and potential silencing via RNA interference; (iii) regulation of the ribosomal machinery that accomplishes translation; and (iv) modulation of the activity state of protein, through, for example, degradation, conformation changes, and phosphorylation. Recalling the circadian clock schematic in Fig. 23.2, the process of

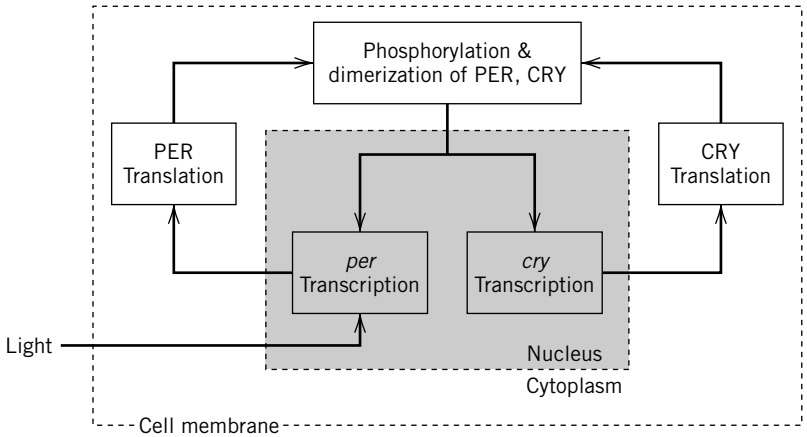


**Figure 23.2** The gene regulatory circuit responsible for mammalian circadian rhythms (by convention, italics and lowercase refer to genes, uppercase refers to proteins).

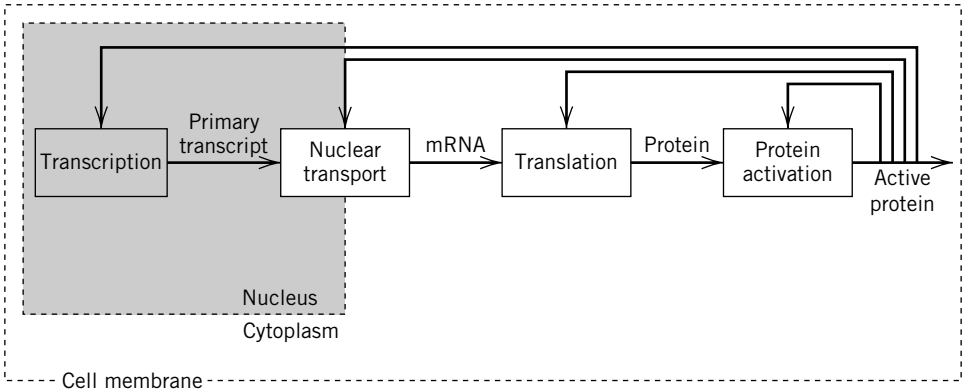


**Figure 23.3** The layers of feedback control in the Central Dogma (modified from Alberts et al., 1998)

controlling the concentration of a phosphorylated form of the PER protein can be broken down into each of the elementary steps indicated by the Central Dogma schematic in Fig. 23.3.

Systems biology holds great promise to revolutionize the practice of medicine, enabling a far more predictive and preventative capability (Hood et al., 2004). As scientists and engineers begin to understand the complex networks of genes and proteins that are regulated through feedback and feedforward control, it is possible to develop novel therapies through systematic modification of these closed-loop systems. These modification sites are referred to as targets, and they are opportunities for the design of drugs by the pharmaceutical sector. A drug may target a particular gene, or a protein, or an activity state of a protein (e.g., phosphorylated form), suggesting that there are multiple intervention points in the Central Dogma process, as depicted in Fig. 23.3. In control terminology, they are potential manipulated variables to restore a healthy state to the network. Likewise, medical scientists and engineers are looking for markers that reveal the pattern of a disease in the signature of the network response. Again, they are understood in control terms as novel sensors that form the basis of an inferential strategy to monitor the status of an unmeasurable disease state. Just as process control engineers test the efficacy of their control system designs through simulation, systems biologists evaluate these new drug targets through extensive simulations of patient populations.

## 23.2 GENE REGULATORY CONTROL

As described in the previous section, genes are regulated through complex feedback control networks. These networks exhibit a remarkable degree of robustness, because the transcription of critical genes is reliable and consistent, even in the face of disturbances from both within the cell and external to the organism. One of the very compelling features of gene regulatory networks is the recurring use of circuit elements that occur in engineering networks. It has been shown that groups of two to four genes exhibit recurring connection topologies, so-called motifs, which have direct analogs in digital electronic circuits (several examples are illustrated in Fig. 23.4). Thus, nature employs these fundamental building blocks in constructing a wide array of gene regulatory networks.

There are a couple of technical terms associated with gene regulatory networks that require explanation. A gene is a portion of the DNA sequence of an organism, which has two primary subregions that are relevant for feedback control: (i) the regulatory or noncoding region can be considered as the input for transcription feedback, and (ii) the coding region determines the products of the expression process, in other words, the output of transcription. The noncoding region can be further divided into discrete regions of separate regulation, called promoters, to which transcription factors bind, leading to activation or inhibition of the expression of the gene (the transcription process). The combination of transcription factors and promoter regions are the controller for the gene transcription process.

There are three dominant network motifs found in *E. coli* (Shen-Orr et al., 2002): (i) a feedforward loop, in which one transcription factor regulates another factor, and, in turn, the pair jointly regulates a third transcript factor; (ii) a single-input multiple-output (SIMO) block architecture; and (iii) a multiple-input multiple-output (MIMO) block architecture, referred to as a densely overlapping regulon by biologists.

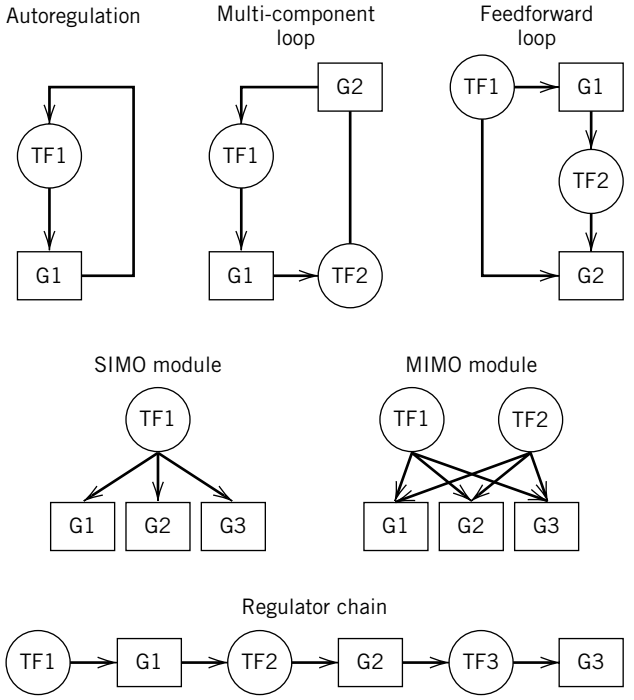


**Figure 23.4** Examples of circuit motifs in yeast (adapted from Lee et al., 2002). The rectangles denote promoter regions on a gene (G1, G2, etc.), and the circles are transcription factors (TF1, TF2, etc.).

A completely different organism, *S. cerevisiae*, has six closely related network motifs (Lee et al., 2002): (i) an autoregulatory motif, in which a regulator binds to the promoter region of its own gene; (ii) a feedforward loop; (iii) a multicomponent loop, consisting of a feedback closed-loop with two or more transcription factors; (iv) a regulator chain, consisting of a cascade of serial transcription factor interactions; (v) a single-input multiple-output (SIMO) module; and (vi) a multiple-input multiple-output (MIMO) module. These motifs are illustrated in Fig. 23.4.

In effect, these studies prove that, in both eukaryotic and prokaryotic systems, cell function is controlled by complex networks of control loops, which are cascading and interconnected with other (transcriptional) control loops. The complex networks that underlie biological regulation appear to be constructed of elementary systems components, not unlike a digital circuit. This lends credibility to the notion that analysis tools from process control are relevant in systems biology.

Some of the analogies between process control concepts and biological control concepts are summarized in Table 23.1, at the level of gene transcription. Keep in mind that there are many levels of analysis in biological circuits, and one can draw comparisons to engineering circuits at each of these levels.

**Table 23.1** Analogies between process control concepts and gene transcription control concepts

| Process Control Concept | Biological Control Analog |
|---|---|
| Sensor | Concentration of a protein |
| Set point | Implicit: equilibrium concentration of protein |
| Controller | Transcription factors |
| Final control element | Transcription apparatus; ribosomal machinery for protein translation |
| Process | Cellular homeostasis |

## EXAMPLE 23.1

The control strategy of gene regulatory circuits can often be approximated using simple logic functions, much like the functions employed in Chapter 18 for batch recipe control. Consider the logic underlying the regulation of the *lacZ* gene, which is involved in sugar metabolism (Ptashne and Gann, 2002). This gene codes for the enzyme β-galactosidase, which is responsible for cleaving lactose, a less efficient source of energy for a bacterium than the preferred glucose supply. The state of the gene (activated or inhibited) is determined by the transcription factors that bind to the regulatory domain of the gene. One of those transcription factors, catabolite activator protein (CAP), binds to the appropriate promoter domain when glucose is absent and lactose is present, leading to the activation of *lacZ*. The other transcription factor, rep (short for Lac repressor), binds to the appropriate promoter domain in the absence of lactose. Once bound, rep inhibits the expression of the gene. If neither rep nor CAP is present, you may assume that only a very small (basal) rate of gene expression occurs.

**(a)** Develop a logic table for the permutations in outcome (transcription of gene *lacZ*) as a function of the two input signals, CAP and rep.

**(b)** Write a simple logic rule for the expression of the *lacZ* gene as a function of the presence of lactose and glucose (ignore the basal state).

### *SOLUTION*

**(a)** The logic table is given in Table 23.2.

**Table 23.2** Logic table for activity state of gene *lacZ* as a function of input signals CAP and rep

| CAP | rep | *lacZ* state |
|---|---|---|
| + | − | off |
| + | + | basal |
| − | + | activated |
| − | − | off |

**(b)** A simple rule for the expression logic is given as:

$lacZ$ = lactose AND (NOT(glucose))

because the gene (and its enzyme product) are only required when the primary sugar source (glucose) is not present and the secondary source (lactose) is present.

### 23.2.1 Circadian Clock Network

Recall from the previous section that the circadian clock orchestrates a number of important metabolic processes in an organism. It does this by regulating the concentration of key proteins in a cycle manner, with a period of (approximately) 24 h. Consider a simplified model of the *Drosophila melanogaster* circadian clock involving two key genes: the period gene (denoted *per*) and the timeless gene (denoted *tim*). Those genes are transcribed into mRNA, exported from the nucleus, and translated into their respective proteins (denoted in Fig. 23.5 by the uppercase convention as PER and TIM). The protein monomers form a dimer, and the dimers of both PER and TIM combine to form a heteromeric complex that reenters the nucleus and suppresses the rate of transcription of the two genes via negative feedback. The kinetic mechanisms for the phosphorylation events are assumed to be Michaelis-Menten form, and the kinetic mechanism for gene regulation (inhibition) follows a Hill mechanism (with a Hill coefficient of 2).

For the assumptions made by Tyson et al. (1999), the two genes can be lumped together, as well as their corresponding proteins and the nuclear and cytoplasmic forms of the dimer. Finally, assuming rapid equilibrium between the monomer and dimer, a second-order set of balances can be developed for the mRNA state $M$ and the protein state $P$. The resulting pair of differential equations captures the dynamics of the feedback-controlled circuit:

$$\frac{dM}{dt} = \frac{v_m}{1 + (P(1-q)/2P_{crit})^2} - k_m M \qquad (23\text{-}1)$$

$$\frac{dP}{dt} = v_p M \frac{k_{p1}Pq + k_{p2}P}{J_p + P} - k_{p3}P \qquad (23\text{-}2)$$

An additional algebraic relationship introduces a more complex dependence of the transcription rate on the protein concentration P:

$$q = \frac{2}{1 + \sqrt{1 + 8K_{eq}P}} \qquad (23\text{-}3)$$

The model parameters and their definitions are a result of the work of Tyson et al. (1999) and are summarized in Table 23.3.

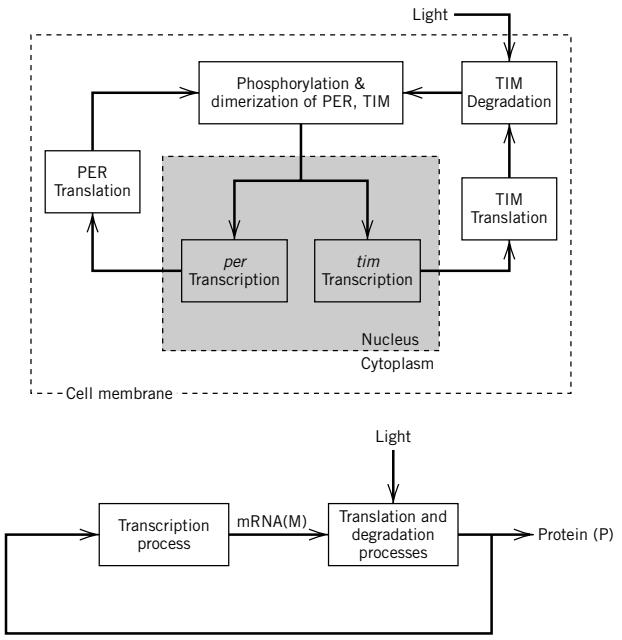


**Figure 23.5** Schematic of negative feedback control of *Drosophila* circadian clock (adapted from Tyson et al., 1999): detailed system (top), and simplified model (bottom).

**Table 23.3** Parameter values for circadian clock circuit in Figure 23.5 ($C_m$ denotes transcript concentration and $C_p$ denotes protein concentration).

| Parameter | Value | Units | Description |
|---|---|---|---|
| $v_m$ | 1 | $C_m h^{-1}$ | Maximum rate of mRNA synthesis |
| $k_m$ | 0.1 | $h^{-1}$ | First-order constant for mRNA degradation |
| $v_p$ | 0.5 | $C_p C_m h^{-1}$ | Rate constant for translation of mRNA |
| $k_{pl}$ | 10 | $C_p h^{-1}$ | $V_{max}$ for monomer phosphorylation |
| $k_{p2}$ | 0.03 | $C_p h^{-1}$ | $V_{max}$ for dimer phosphorylation |
| $k_{p3}$ | 0.1 | $h^{-1}$ | First-order rate constant for proteolysis |
| $k_{eq}$ | 200 | $C_p^{-1}$ | Equilibrium constant for dimerization |
| $P_{crit}$ | 0.1 | $C_p$ | Dimer concentration at half-maximum transcription rate |
| $J_p$ | 0.05 | $C_p$ | Michaelis constant for protein kinase |

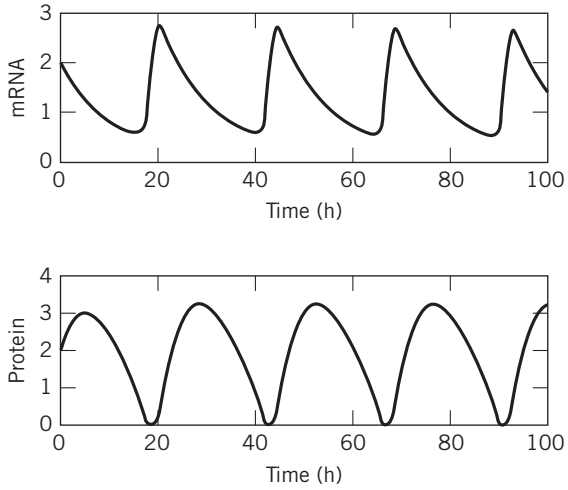


**Figure 23.6** Simulation of the circadian clock model.

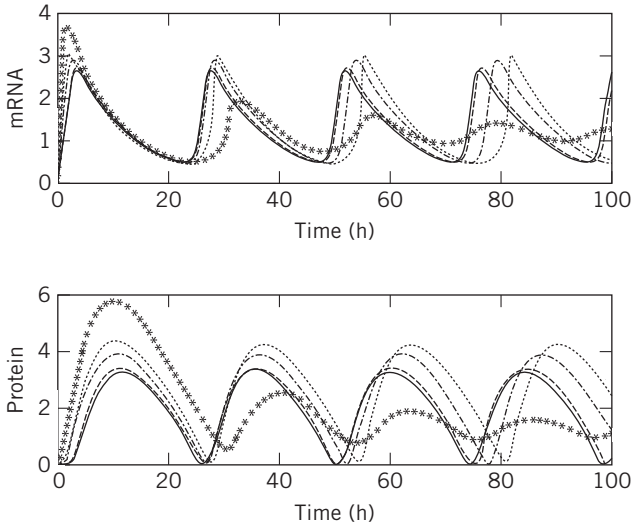


**Figure 23.7** Simulation of circadian clock model for varying values of $\nu_m$ [1.0 (solid), 1.1 (dashed), 1.5 (dash-dot), 2.0 (dotted), 4.0 (asterisk)].

Using a computer package, such as Simulink/MATLAB, the gene regulatory circuit using these defined parameters can be simulated with initial values of $M$ and $P$ equal to [2.0; 2.0]. A 100-h simulation is shown in Fig. 23.6; the period can be calculated from either the mRNA (M) or the Protein (P) trajectory (e.g., time between peaks) and is 23.2 h (i.e., approximately 24 h or "circadian").

A common property of biological closed-loop circuits is that they exhibit remarkable robustness to disturbances and fluctuations in operating conditions. For example, the clock should maintain a nearly 24-hr period, even though the organism is exposed to temperature changes, which affect the rates of biochemical reactions. The model circadian clock can be simulated by perturbing values of the kinetic constants. The same clock simulation is evaluated for the following values of the parameter $\mu_m$: [1.0; 1.1; 1.5; 2.0; 4.0]. The period of the clock lengthens as $\mu_m$ is increased, as shown in Fig. 23.7. The period increases as follows: [23.2; 23.5 25.5; 26.4] corresponding to the first four values of $\mu_m$. At the extreme value of 4.0, oscillations are no longer observed, and the system settles to a stable equilibrium. The stability of the oscillations is quite remarkable for such large perturbations in $\mu_m$ (over 100%).

Another important feature of the circadian clock is its ability to entrain (i.e., track) an external signal (sunlight), so that the period of the oscillations of mRNA and Protein match exactly the period of the external signal. In this manner, the organism's clock is reset to a period of precisely 24 h. Tyson et al. (1999) show that this can be simulated in the present model by switching the value of $K_{eq}$ to emulate dark–light cycles (i.e., using a square wave with even intervals of light and dark in a 24-h period). In the fly, sunlight appears to modulate the rate of degradation of one of the key proteins in the circuit. This can be achieved in the same simulation model by altering $K_{eq}$, between 100 and 200, and observing the period of the driven system. Fig. 23.8 illustrates that the oscillations in mRNA and Protein do indeed exhibit a period equal to the forcing signal (in this case, 20 h).

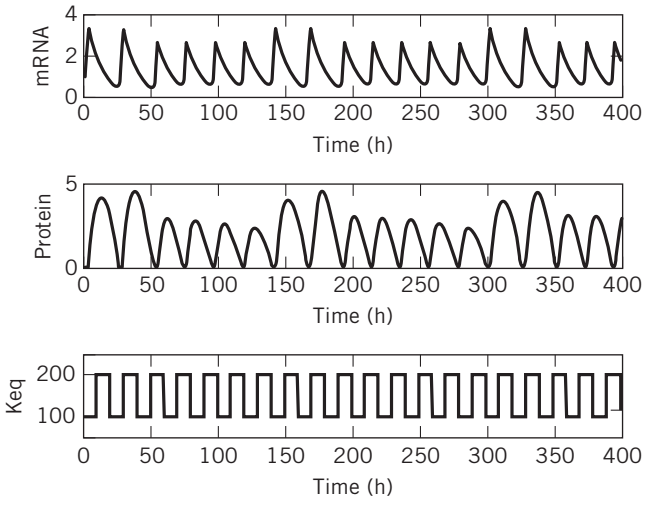


**Figure 23.8** Simulation of circadian clock model for entraining signal with period of 20 h.

## 23.3 SIGNAL TRANSDUCTION NETWORKS

The gene regulatory networks of the previous section are often activated by cues or signals that originate from outside the cell. This is of tremendous importance for unicellular organisms that must sense the environment for survival, but it is also of critical importance for multicellular organisms that require robust coordinated behavior from, for example, a group of cells that constitute a tissue or an organ. A particularly relevant set of such cues are *ligands* (from the Latin "to bind"), which are molecules that bind to proteins that typically span the surface membrane of a cell. These ligands, called receptors, induce particular responses within the cell, depending on the conditions. They include a number of interesting stimulus-response mechanisms (Lauffenburger and Linderman, 1993):

- Growth factors → cell division
- Necrosis factor → programmed cell death (apoptosis)
- Chemoattractant → chemotaxis
- Insulin → glucose uptake
- Neurotransmitter → secretion by nerve cell
- Extracellular matrix (ECM) protein → adhesion

Once the ligand binds to the receptor, it initiates a series of biochemical reactions that induce a short-term response (e.g., phosphorylation state of an intermediate protein) and/or a longer-term response as a result of a regulated gene response. These networks respond relatively rapidly, exhibiting dynamics with characteristic time scales of seconds to minutes. A cell is often presented with multiple, competing cues, and it processes that information in rich signal transduction networks, to result in the appropriate cellular fate, depending on the context.

In this section, we highlight several signal transduction cascades, to illustrate the rich processing dynamics manifested by these networks.

### 23.3.1 Chemotaxis

The process of chemotaxis is the directed motion of a cell or cellular organism toward a chemical source, typically a food molecule. This mechanism is also invoked in the response to a detected toxin (i.e., motion away from that source) and is involved in more complex processes, such as development. The process is initiated by the detection of a ligand (e.g., a food molecule) at the cell surface, which invokes a signal transduction cascade and results in the alteration of the motor apparatus responsible for moving the cell.

A simplified version of the biochemical pathway that underlies chemotaxis in *E. coli* is shown in Fig. 23.9. The binding of an attractant molecule (ligand) to the receptor complex CheW-CheA (denoted as W-A) induces the phosphorylation of protein CheY (Y), and the phosphorylated form (Yp) invokes a tumbling motion from the bacteria's flagella. This tumbling motion allows the organism to reorient and search the surrounding space; otherwise, the organism proceeds in a straight run. The ability of CheW-CheA to phosphorylate CheY depends on the methylation state of that complex, which is fine-tuned by the proteins CheR (R) and the phosphorylated form of CheB (Bp), as illustrated in the figure. Feedback is evident in Fig. 23.9, because CheB phosphorylation is mediated by the CheW-CheA complex.

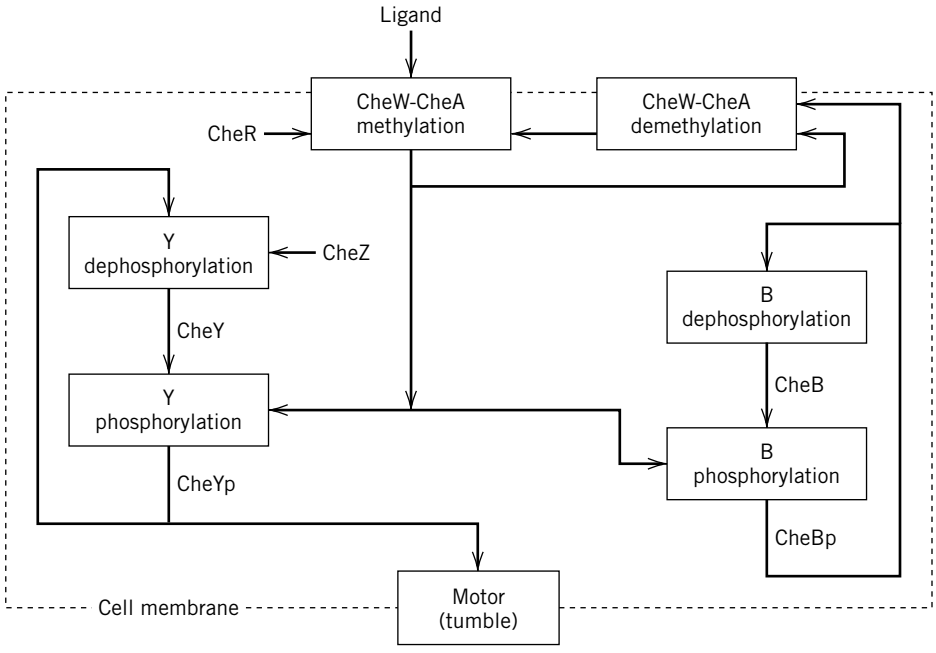


**Figure 23.9** Schematic of chemotaxis signaling pathway in *E. coli* (adapted from Rao et al., 2004).

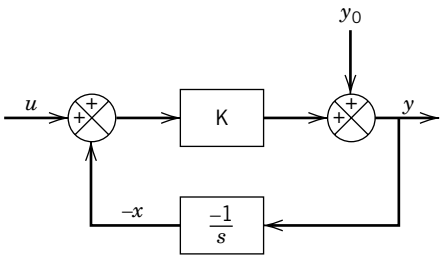


**Figure 23.10** Integral control feedback circuit representation of chemotaxis (adapted from Yi et al., 2000).

The signal transduction system that mediates chemotaxis exhibits a type of adaptation in which the response to a persistent stimulus is reset to the pre-stimulus value, thereby enabling an enhanced sensitivity. Several mechanistic explanations can be postulated for this robust behavior, including the following: (i) precise fine-tuning of several parameters to yield a consistent (robust) response under varied conditions, or (ii) inherent regulation that yielded this robust behavior. Utilizing process control principles, it has been demonstrated that the regulatory system exploits integral feedback control to achieve the robust level of adaptation exhibited in chemotaxis (Yi et al., 2000). The chemotaxis network can be reduced to the simple block diagram in Fig. 23.10, in which $u$ denotes the chemoattractant, $y$ denotes the receptor activity, and $-x$ denotes the methylation level of the receptors. It is left as an exercise to show that this circuit ensures that perfect adaptation is achieved (i.e., the receptor activity always resets to zero asymptotically).

This understanding suggests that many seemingly complex biological networks may employ redundancy and other structural motifs or modules to achieve relatively simple overall system behavior.

### 23.3.2 Insulin-Mediated Glucose Uptake

Muscle, liver, and fat cells in the human body take up glucose as an energy source in response to, among other signals, the hormone insulin, which is secreted by the pancreas. As discussed in Chapter 22, the release of insulin is regulated in a feedback manner by the blood glucose level. In Type 2 diabetes, the insulin signal transduction network is impaired such that insulin does not lead to glucose uptake in these cells. A simplified model of the insulin signaling network can be decomposed into three submodules, as shown in Fig. 23.11. The first submodule describes insulin receptor dynamics: insulin binds to insulin receptor, causing subsequent receptor autophosphorylation. The receptor can also be recycled, introducing additional dynamics in the network. The second submodule describes the phosphorylation cascade downstream from the insulin receptor. The final submodule describes the activation of movement and fusion of specialized glucose transporter

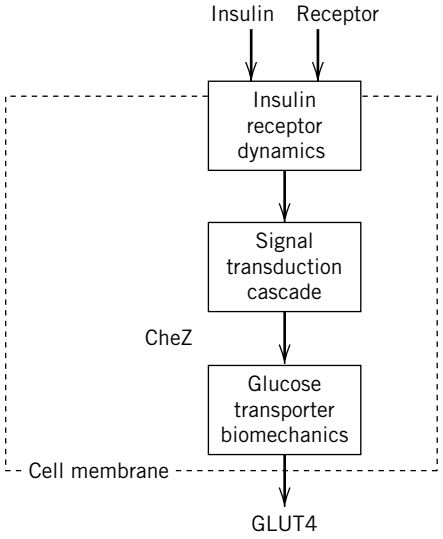


**Figure 23.11** Simplified insulin signaling pathway for glucose uptake.

(GLUT4) storage vesicles with the plasma membrane by the intermediate proteins from the second module. These GLUT4 transporters allow glucose molecules to enter the cell. Each of the three modules contains submodules that consist of layers of feedback.

### 23.3.3 Simple Phosphorylation Transduction Cascade

In signal transduction, a receptor signal is processed in a cascaded pathway, to yield a cellular response. For the example considered here, the processing consists of a sequence of kinase- and phosphatase-catalyzed reaction steps, consisting of phosphorylation and dephosphorylation, respectively. The key performance attributes of such a system are (i) the speed at which a signal arrives to the destination, (ii) the duration of the signal, and (iii) the strength of the signal. Under conditions of weak activation (low degree of phosphorylation), the individual steps in the signal transduction cascade can be modeled as a set of linear ODEs (Heinrich et al., 2002):

$$\frac{dX_i}{dt} = \alpha_i X_{i-1} - \beta_i X_i \qquad (23\text{-}4)$$

where $\alpha_i$ is a pseudo first-order rate constant for phosphorylation, $\beta_i$ is the rate constant for dephosphorylation, and $X_i$ is the phosphorylated form of the kinase (i). Assume that the cascade consists of four stages (levels of phosphorylation), that the corresponding rate constants are equal for all stages ($\alpha_i = \alpha$; $\beta_i = \beta$), and that the receptor inactivation is approximated as an exponential decay with time constant $1/\lambda$ (see Fig. 23.12). The resulting cellular response can be written in the Laplace domain as,

$$Y(s) = \left(\frac{s}{\frac{1}{\lambda}s + 1}\right)\left(\frac{\alpha^4}{(s+\beta)^4}\right)R(s) \qquad (23\text{-}5)$$

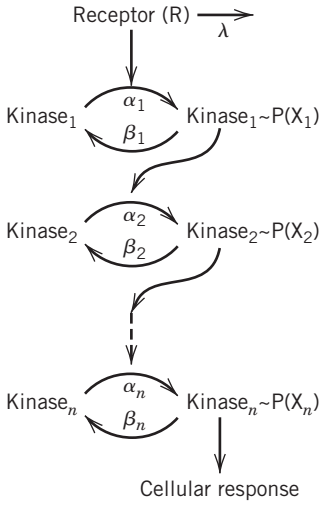


**Figure 23.12** Schematic of fourth-order signal transduction cascade for Example 23.3, combined with first-order receptor activation (adapted from Heinrich et al., 2002).

where *R(s)* is the receptor input and *Y(s)* is the cellular response.

If the signaling $T_{sig}$ time is defined as the average time to activate a kinase, a suitable expression in the time domain for this quantity is:

$$T_{sig} = \frac{\int_0^\infty ty(t)dt}{\int_0^\infty y(t)dt} \tag{23-6}$$

where *y(t)* is the unit step response (*R(s)* = 1/s in (23-5)). It is possible to derive the analytical expression for the signaling time for this network. Recalling a few rules from Laplace transforms (see Appendix A):

$$\mathcal{L}(tf(t)) = -\frac{d}{ds}F(s) \tag{23-7}$$

and

$$\mathcal{L}\left(\int_0^\infty f(t)dt\right) = F(s=0) \tag{23-8}$$

Then the following expression for the signal time can be derived:

$$T_{sig} = \frac{\left(-\frac{d}{ds}Y(s)\right)_{s=0}}{Y(0)}$$

$$= \frac{[\lambda\alpha^4(s+\lambda)^{-2}(s+\beta)^{-4} + 4\lambda\alpha^4(s+\lambda)^{-1}(s+\beta)^{-5}]_{s=0}}{\alpha^4/\beta^4} \tag{23-9}$$

which simplifies to:

$$T_{sig} = \frac{1}{\lambda} + \frac{4}{\beta} \tag{23-10}$$

Notice that the average time through the network (i.e., the signaling time) is not dependent on the rate of phosphorylation ($\alpha$).

## SUMMARY

In this chapter, a number of biological circuit diagrams have been introduced that illustrate the rich array of dynamics and feedback control that exist in all living organisms. Two particular biological processes were considered: the regulation of gene transcription and the protein signal transduction that characterizes cellular stimulus-response mechanisms. The recurring motifs of feedback and feedforward control motivated the application of process control analysis to these problems, to shed light on both the healthy functioning state as well as to promote the investigation of therapies for cases where the natural circuit is impaired (i.e., a disease state).

The rapidly developing field of systems biology continues to make great advancements in the area of medical problems, and the increased understanding of the biological circuits underlying diseases will likely lead to novel therapeutic strategies, as well to the discovery of new drugs. More information is available in more specialized books, including those of Klipp et al. (2005), Palsson (2006), and Alon (2007).

## GLOSSARY

***Eukaryote:*** an organism that is comprised of cells (or possibly a single cell, as in yeast) that are divided into substructures by membranes, notably containing a nucleus. Examples include animals, plants, and fungi.

***Kinase:*** an enzyme that catalyzes the transfer of phosphate group to a substrate, leading to phosphorylation of that substrate.

***Prokaryote:*** an organism that is comprised of a single cell that does not contain a separate nucleus. Examples include bacteria and archae.

***Promoter:*** a region of a DNA involved in the regulation of transcription of the corresponding gene.

## REFERENCES


Alon, U., *An Introduction to Systems Biology: Design Principles of Biological Circuits*, Chapman & Hall/CRC, New York, 2007.

Alberts, B., D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walters. *Essential Cell Biology*, Garland Pub., Inc., New York, 1998.

Barabasi, A. L., Network Biology: Understanding the Cell's Functional Organization, *Nature Rev. Genetics*, **5**, 101 (2004).

Boulos, Z., M. M. Macchi, M. P. Sturchler, K. T. Stewart, G. C. Brainard, A. Suhner, G. Wallace, and R. Steffen, Light Visor Treatment for Jet Lag after Westward Travel across Six Time Zones, *Aviat. Space Environ. Med*., **73**, 953 (2002).

Campbell, A. M., and L. J. Heyer. *Discovering Genomics, Proteomics, & Bioinformatics*, Benjamin Cummings, San Francisco, CA, 2007.

Daan, S., and C. S. Pittendrigh, A Functional Analysis of Circadian Pacemakers in Nocturnal Rodents. II. The Variability of Phase Response Curves, *J. Comp. Physiol*., **106**, 253 (1976).

Dunlap, J. C., J. J. Loros, and P. J. DeCoursey (Eds.). *Chronobiology: Biological Timekeeping*, Sinauer Associates, Inc., Sunderland, MA, 2004.

El-Samad, H., S. Prajna, A. Papachristodoulou, J. Doyle, and M. Khammash, Advanced Methods and Algorithms for Biological Networks Analysis, *Proc. IEEE*, **94**, 832 (2006).

Heinrich, R., B. G. Neel, and T. A. Rapoport, Mathematical Models of Protein Kinase Signal Transduction, *Molecular Cell, **9**, 957 (2002).*

Herzog, E. D., S. J. Aton, R. Numano, Y. Sakaki, and H. Tei, Temporal Precision in the Mammalian Circadian System: A Reliable Clock from Less Reliable Neurons, *J. Biol. Rhythms*, **19**, 35 (2004).

Hood, L., J. R. Heath, M. E. Phelps, and B. Lin, Systems Biology and New Technologies Enable Predictive and Preventative Medicine, *Science*, **306**, 640 (2004).

Kitano, H., Systems Biology: A Brief Overview, *Science*, **295**, 1662 (2002).

Klipp, E., R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice*, Wiley-VCH, Weinheim, 2005.

Lauffenburger, D. A., and J. J. Linderman. *Receptors: Models for Binding, Trafficking, and Signaling*, Oxford University Press, New York, 1993.

Lee, T. I., N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford and R. A. Young, Transcriptional Regulatory Networks in *Saccharomyces cerevisiae*. *Science*, **298**, 799 (2002).

Liu, A. C., D. K. Welsh, C. H. Ko, H. G. Tran, E. E. Zhang, A. A. Priest, E. D. Buhr, O. Singer, K. Meeker, I. M. Verma, F. J. Doyle III, J. S. Takahashi, and S.A. Kay, Intercellular Coupling Confers Robustness against Mutations in the SCN Circadian Clock Network, *Cell*, **129**, 605 (2007).

National Research Council, *Network Science*, National Academies Press, Washington DC, 2005.

Palsson, B., *Systems Biology: Properties of Reconstructed Networks*, Cambridge University Press, New York, 2006

Ptashne, M. and A. Gann. *Genes and Signals*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2002.

Rao, C.V., J. R. Kirby, and A. P. Arkin, Design and Diversity in Bacterial Chemotaxis: A Comparative Study in *Escherichia coli* and *Bacillus subtilis*, *PLoS Biology*, **2**, 239 (2004).

Reppert, S. M. and D. R. Weaver, Coordination of Circadian Timing in Mammals, *Nature*, **418**, 935 (2002).

Sedaghat, A. R., A. Sherman, and M. J. Quon, A Mathematical Model of Metabolic Insulin Signaling Pathways, *Amer. J. Physio-End. Met*., **283**, E1084 (2002).

Shen-Orr, S. S., R. Milo, S. Mangan and U. Alon, Network Motifs in the Transcriptional Regulation Network of *Escherichia coli*, *Nature Genetics*, **31**, 64 (2002).

Tyson, J. J., C. I. Hong, C. D. Thron, and B. Novak, A Simple Model of Circadian Rhythms Based on Dimerization and Proteolysis of PER and TIM, *Biophys. J*., **77**, 2411 (1999).

Yi, T. M., Y. Huang, M. I. Simon, and J. Doyle, Robust Perfect Adaptation in Bacterial Chemotaxis through Integral Feedback Control, *Proc. Nat. Acad. Sci. USA*, **97**, 4649 (2000).


## EXERCISES

**23.1** In this exercise, treat the components as simple (reactive) chemical species and perform the appropriate (dynamic) material balance. Assume that a messenger RNA (*mRNA*) is produced by a constant (basal) expression rate from a particular gene. In addition, assume that the *mRNA* degrades according to a first-order decay rate.

**(a)** Write the equation for the dynamics of the mRNA concentration as a function of the expression rate ($G_0$) and the decay rate constant ($k_d^{mRNA}$).

**(b)** Assume that each *mRNA* molecule is translated to form *p* copies of a protein product, *P*. Furthermore, the protein is subject to first-order degradation, with a decay rate constant ($k_P^{mRNA}$). Write the equation for the dynamics of the protein concentration.

**(c)** Assume that the system has been operating for some time at a constant gene expression rate ($G_0$), and then the expression rate changes instantaneously to a value $G_1$. Derive an analytical expression for the transient responses for *mRNA* and *P*.

**23.2** Consider the block diagram in Fig. E23.2 of the multiple feedback loops involved in the Central Dogma schematic from Fig. 23.3, namely genetic regulation ($C_1$), translational regulation ($C_2$), and enzyme inhibition ($C_3$). Assume that the processes $P_1$, $P_2$, and $P_3$ obey first-order dynamics, with corresponding gains and time constants ($K_i$, $\tau_i$).

**(i)** Derive the transfer function from the external input (*u*) to the output (*y*) for each of the three cases shown in Figure E23.2 (a), (b), (c).

**(ii)** Assume that the feedback mechanisms operate via proportional control with corresponding controller gains ($K_{ci}$). Derive the closed-loop transfer function from the external input (*u*) to the output (*y*) in block diagram (b).

**(iii)** Consider a simplified biological circuit in which only genetic regulation is active ($C_1$). Derive the closed-loop transfer function and comment on the key differences between this transfer function and the one from part (b).

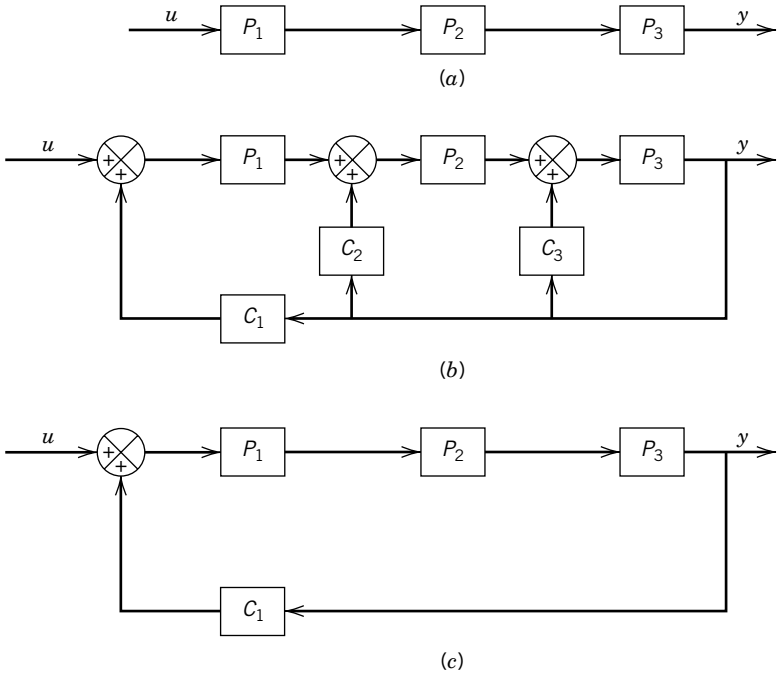


**Figure E23.2**

**(iv)** Give several reasons why the natural feedback architecture with all three controllers operating is more effective than the control architecture in part (c).

**23.3** As a specific biological example for Exercise 23.2 and Figure E23.2(b),* the synthesis of tryptophan can be described by the following set of material balances:

$$\frac{d}{dt}[O_R] = k_1[O_t]C_1[T] - k_{d1}[O_R] - \mu[O_R]$$

$$\frac{d}{dt}[mRNA] = k_2[O_R]C_2[T] - k_{d2}[mRNA] - \mu[mRNA]$$

$$\frac{d}{dt}[E] = k_3[mRNA] - \mu[E]$$

$$\frac{d}{dt}[T] = k_4C_3[T][E] - g\frac{[T]}{[T]+K_g} - \mu[T]$$

where $k_1$, $k_2$, $k_3$, and $k_4$ represent kinetic rate constants for the synthesis of free operator, mRNA transcription, translation, and tryptophan synthesis, respectively. Parameters $O_t$, $\mu$, $k_{d1}$, and $k_{d2}$ refer to total operator site concentration, specific growth rate of *E. coli*, degradation rate constants of free operator $O_R$, and mRNA, respectively. $E$ and $T$ represent concentrations of enzyme anthranilate synthase and tryptophan, respectively, in the cell. $K_g$ and $g$ are the half saturation constant and kinetic constant for the uptake of tryptophan for protein synthesis in the cell. Model parameter values are as follows: $k_1 = 50$ min$^{-1}$; $k_2 = 15$ min$^{-1}$; $k_3 = 90$ min$^{-1}$; $k_4 = 59$ min$^{-1}$; $O_t = 3.32$ nM; $k_{d1} = 0.5$ min$^{-1}$; $k_{d2} = 15$ min$^{-1}$; $\mu = 0.01$ min$^{-1}$; $g = 25$ $\mu$M. min$^{-1}$; $K_g = 0.2$ $\mu$M. Here, controllers $C_1(T)$, $C_2(T)$, and $C_3(T)$ represent repression, attenuation, and inhibition, respectively, by tryptophan and are modeled by a particular form of Michaelis-Menten kinetics (the Hill equation) as follows:

$$C_1(T) = \frac{K_{i,1}^{\eta_H}}{K_{i,1}^{\eta_H} + T^{\eta_H}},\ C_2(T) = \frac{K_{i,2}^{1.72}}{K_{i,2}^{1.72} + T^{1.72}},\ C_3(T) = \frac{K_{i,3}^{1.2}}{K_{i,3}^{1.2} + T^{1.2}}$$

$K_{i,1}$, $K_{i,2}$, and $K_{i,3}$ represent the half-saturation constants, with values $K_{i,1} = 3.53$ $\mu$M; $K_{i,2} = 0.04$ $\mu$M; $K_{i,3} = 810$ $\mu$M, whereas sensitivity of genetic regulation to tryptophan concentration, $\eta_H = 1.92$.

**(a)** Draw a block diagram, using one block for each of the four states. Comment on the similarities between this diagram and schematic (b) in Fig. E23.2.

**(b)** Simulate the response of the system to a step change in the concentration of the medium (change $g$ from 25 to 0 $\mu$M).

**(c)** Calculate the rise time, overshoot, decay ratio, and settling time for the closed-loop response.

**(d)** Omit the inner two feedback loops (by setting $C_2$ and $C_3$ to 0) and change the following rate constants: $K_{i,1} = 8 \times 10^{-8}$ $\mu$M; $\eta_H = 0.5$. Repeat the simulation described in part (b), and obtain the new closed-loop properties for this network (compared to part (c)).

**23.4** Consider Section 23.3.3, where the dynamic properties of a signal transduction were analyzed. Two properties of interest are the signal duration and the amplitude of the signal.

**(a)** The following definition is used for signal duration:

$$T_{dur} = \sqrt{\frac{\int_{t=0}^{\infty} t^2 y(t)dt}{\int_{t=0}^{\infty} y(t)dt} - T_{sig}^2}$$

where $T_{sig}$ was defined in Section 23.3.3. Use Laplace transforms to derive an expression for the signal duration as a function of the parameters in the phosphorylation cascade.

*The authors acknowledge Profs. Bhartiya, Venkatesh, and Gayen for their help with formulating this problem.

**(b)** Define the signal amplitude as:

$$A = \frac{\int_{t=0}^{\infty} y(t)dt}{2T_{dur}}$$

Use Laplace transforms to derive an expression for the signal amplitude as a function of the parameters in the phosphorylation cascade.

**23.5** Consider the simplified version of the chemotaxis circuit in Fig. 23.10.

**(a)** Derive the conditions for the process gain $K$ that ensure that the receptor activity is always reset to zero and even for the case of a persistent ligand signal.
**(b)** Show that the closed-loop transfer function from the ligand to the receptor activity is equivalent to a first-order transfer function with numerator dynamics.
**(c)** Comment on the biological relevance of the result in part (b), particularly for a ligand signal that is fluctuating.

**23.6** An interesting motif in biological circuits is a switch, in which the system can change from (effectively) one binary state to another. An analysis of a continuous reaction network reveals a rise to a switchlike response (also referred to as ultrasensitivity). Consider interconversion of a protein from its native state $P$ to an activated form $P^*$, catalyzed by the enzymes $E_1$ and $E_2$:

$$P + E_1 \leftrightarrow PE_1 \leftrightarrow P^* + E_1 \leftrightarrow P^* + E_2 \rightarrow P^*E_2 \rightarrow P + E_2$$

**(a)** Assume that all reaction steps obey mass-action kinetics. What is the steady-state dependence of $P^*$ as a function of the concentration of $E_1$? (Assume that total amount of $E_1$ $E_2$, and $P$ are all constant and that $P$ is in excess compared to $E_1$ and $E_2$.)
**(b)** Alternate starting point for problem: you should be able to rearrange the solution as follows:

$$\frac{V_1}{V_2} = \frac{P^*/P_T(1 - P^*/P_T + K_1)}{(1 - P^*/P_T)(P^*/P_T + K_2)}$$

where $V_1$ is proportional to the total $E_1$ in the system, $E_2$ is proportional to the total $V_2$ in the system, $P_T$ is the total protein concentration (in all forms), and $K_1$ and $K_2$ are suitable combinations of the rate constants for the reactions previously described.

For $K_1 = 1.0$, $K_2 = 1.0$, plot the steady-state locus of solutions for $(P^*/P_T)$ versus $V_1/V_2$.

**(c)** Assume that the two enzymes operate in a saturated regime, i.e., the reactions follow zero-order kinetics with respect to the enzymes. Use the expression from part (b) to plot the steady-state locus for this extreme situation (i.e., $K_1 = 0$, $K_2 = 0$).

**(d)** Comment on the difference in shape of the gain functions in parts (b) and (c). Based on the initial problem description, explain how biology can produce switchlike behavior in this system.