Neuron, Vol. 26, 567–580, June, 2000, Copyright ©2000 by Cell Press

Mathematical Modeling of Gene Networks Review

Paul Smolen, Douglas A. Baxter, and John H. Byrne*
Department of Neurobiology and Anatomy
W.M. Keck Center for the Neurobiology of Learning and Memory
The University of Texas–Houston Medical School
Houston, Texas 77225

The expression of clusters of genes can be regulated by stimulus-induced activation of signal transduction pathways that modulate the activity of transcription factors (TFs). TFs regulate transcription by binding to stretches of DNA known as responsive elements (reviewed by Karin, 1994). Many genes are regulated by multiple TFs. Furthermore, genes coding for TFs are often activated or repressed by their own products as well as by other TFs. A large number—on the order of thousands—of such interactions together constitute the total array of genetic regulation in neurons and other cells. The term "gene network" can be used to refer to clusters of genes that satisfy two conditions. The expression of the genes in the cluster is affected in concert by specific stimuli, such as exposure to a hormone or neurotransmitter. Also, the protein products of some members of the cluster act as TFs to regulate the expression of other members.

An example of a gene network, which has been repeatedly implicated in neuronal plasticity and in learning and memory, is the cluster of genes that code for the Ca^{2+}/cAMP-responsive element binding protein (CREB) family of TFs (Tully et al., 1994; O'Leary et al., 1995; Sassone-Corsi; 1995; Yin and Tully, 1996; Bartsch et al., 1998; Finkbeiner and Greenberg, 1998; Glazewski et al., 1999; Mayford and Kandel, 1999). Figure 1 illustrates components of this network. Stimuli such as binding of hormones or neurotransmitters to receptors, or Ca^{2+} influx, activate kinases that in turn phosphorylate and activate TFs of the CREB family. These TFs can then bind to Ca^{2+}/cAMP-responsive elements (CREs) to activate genes crucial for neuronal plasticity. Some of these TFs, such as CREB2, can act as transcriptional repressors. Within this network some gene products regulate the transcription of other genes. A negative-feedback loop relies on a TF termed inducible Ca^{2+}/cAMP-responsive early repressor (ICER) (Sassone-Corsi, 1995). ICER is induced by CREB. ICER then binds to CREs and represses its own transcription as well as that of CREB. Also, a positive-feedback loop appears to exist, with CREB binding to CREs and activating its own transcription (Sassone-Corsi, 1995; Walker et al., 1995).

To understand and explore the behavior of complex biochemical systems such as the gene network in Figure 1, it is helpful to use the conceptual framework of mathematical modeling. Intuition alone cannot comprehend the effects of the multiple, simultaneous regulatory interactions within these systems. Models are particularly well suited for predicting the effects of nonlinear interactions such as those dependent on oligomerization of TFs, and for predicting the effects of biochemical time delays, such as are required for intracellular transport of macromolecules. Thus, for a specific gene network, a mathematical model integrates a variety of molecular features into a coherent picture of network operation. Under appropriate conditions, such a model could help to predict responses of normal and mutant organisms to stimuli, such as the efficacy of formation of long-term memory (LTM) and the timing and phase shifts of circadian oscillations. In the gene network of Figure 1, the negative and positive feedback loops discussed above might support complex dynamics such as oscillations in transcription rate. With the help of mathematical modeling, one can predict whether the network structure and the values of system parameters would generate complex dynamics under particular experimental conditions.

Mathematical models are also useful for integrating recent, large gene expression datasets, such as time courses of gene expression in response to pharmacological or hormonal stimuli, or during development (Wen et al., 1998). The advent of "DNA chips" (Gerhold et al., 1999) allows the simultaneous measurement of expression time courses for up to ~10,000 genes. Mathematical modeling and analysis is essential to extract meaningful information concerning genetic regulation from such large, complex datasets. Many genes have been identified only recently, as the genomes of a variety of organisms are characterized (e.g., *Drosophila*, Adams et al., 2000; *C. elegans*, The *C. elegans* Sequencing Consortium, 1998; and man, Collins et al., 1998). Large gene expression datasets will need to be analyzed to help determine the function and regulation of these new genes.

Many studies have examined the behavior of generic gene network models incorporating common mechanisms of regulation by TFs, and these models have proven quite valuable. In this review, we discuss examples that illustrate the value of such models (see also Smolen et al., 2000). The discussion is limited to regulation of transcription by TFs, because comparatively few studies have modeled other forms of regulation. New experimental techniques that provide expression time courses for large sets of genes are also briefly discussed, as are mathematical techniques that help to analyze these time courses to identify regulatory interactions. To familiarize the reader with modeling techniques, we begin by comparing the advantages and disadvantages of commonly used methods for modeling gene networks.

Techniques for Gene Network Modeling Focus on Different Levels of Detail

There are two basic approaches in use for gene network modeling—the "logical-network" or "Boolean" method, and the "dynamic-systems" method using ordinary differential equations.

In the logical-network method (Somogyi and Sniegoski, 1996), a relatively simple level of detail is used. The expression of each gene in the network is assumed

*To whom correspondence should be addressed (e-mail: jbyrne@nba19.med.uth.tmc.edu).

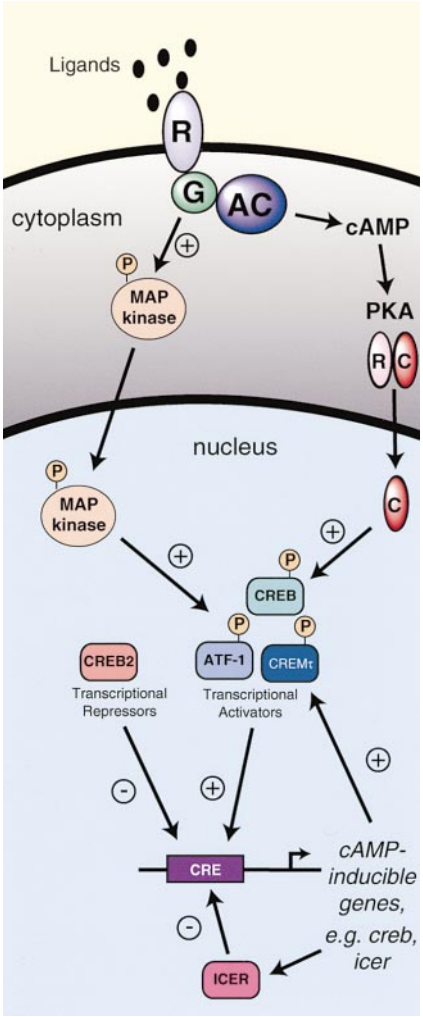

Figure 1. Schematic of the Gene Network that Codes for CREB and Related TFs

Ligands (e.g., hormones or neurotransmitters) bind to receptors in the plasma membrane and activate G proteins. Adenylyl cyclase (AC) is thereby stimulated, and the resultant elevation of cAMP activates PKA, shown as a dimer of regulatory (R) and catalytic (C) subunits. Ca^{2+} influx can also lead to activation of AC, as well as of CaM kinases. MAP kinases may also be activated. Activation of any of these kinases can lead to phosphorylation of CREB and related TFs. These TFs then activate the transcription of genes with Ca^{2+}/cAMP-responsive elements in their promotor regions. Possible feedback interactions among genes of the CREB family are shown. Autoinduction of CREB synthesis may provide positive feedback, and induction of the repressor ICER may provide negative feedback.

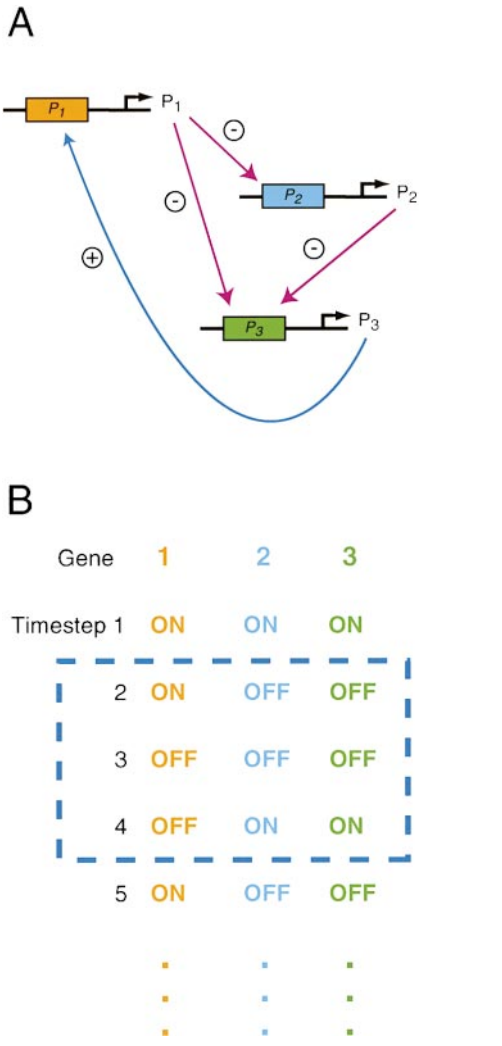

Figure 2. A Simple Gene Network Model Regarding Genes as ON–OFF Switches

(A) A network of three genes. The product of gene 1 inhibits transcription of genes 2 and 3, the product of gene 2 also inhibits gene 3, and the product of gene 3 activates gene 1.

(B) A simulation of the network of (A) using the "logical-network" or "Boolean" modeling technique. Irrespective of the initial state of the network, it quickly settles into a repeating cycle of three states (outlined by dashed box).

to be either ON or OFF—no "intermediate" activity levels are considered. Regulation of some network genes by the products of other genes is characterized by simple ON–OFF relations. An example is given in Figure 2A. Here, a simple network of three genes is shown. The product of gene 1 inhibits transcription of genes 2 and 3, the product of gene 2 also inhibits gene 3, and the product of gene 3 activates gene 1. Genes 2 and 3 have a basal transcription rate in the absence of inhibition, whereas gene 1 is not transcribed in the absence of activation. The corresponding ON–OFF relation for gene 1 is that gene 1 is only ON at a given time step of a computer simulation if gene 3 was ON at the previous time step. Similarly, gene 3 is ON only if genes 1 and 2 were OFF at the previous time step. Such simple rules allow large time steps to be used and minimize the computational time required for simulations. Figure 2B illustrates a simulation with this simple network. Irrespective of the initial state of the network (e.g., all genes ON in Figure 2B), the network quickly settles into a repeating cycle of three states (outlined by dashed box in Figure 2B). The ON–OFF regulatory relations determine this behavior.

As an example of qualitative insights obtained by this method, Thomas et al. (1995) illustrated that positive-feedback loops are important for allowing multiple steady states of gene expression rates, and that negative-feedback loops are important for maintaining homeostasis of gene expression rates. However, as might

be expected with a method that idealizes genes as ON–OFF switches, there are considerable limitations. For example, stable oscillations produced by a logical-network model may not be observed in a more accurate model that describes gene expression rates as real variables that can take on a range of values (Bagley and Glass, 1996; Mestl et al., 1996). A steady state of a logical-network model may also not correspond to a steady state in a model that includes gene expression rates as real variables (Glass and Kauffman, 1973).

The dynamic-systems approach uses ordinary differential equations to describe the rates of change of the concentrations of gene products—mRNAs and proteins. Terms in these differential equations describe how gene expression rates are modified by changes in the levels of transcription factors or other effector molecules (Smith, 1987a; Mahaffey et al., 1992; Mestl et al., 1996; Leloup and Goldbeter, 1998; Smolen et al., 2000). For even greater detail, one can incorporate stochastic fluctuations in molecule numbers due to randomness in the times of synthesis or degradation of individual macromolecules (Arkin et al., 1998; Smolen et al., 2000). Mathematical tools developed for the dynamic-systems approach can provide a detailed understanding of the different types of nonlinear behaviors exhibited by gene networks. A technique termed bifurcation analysis (Guckenheimer and Holmes, 1983) can be used to determine ranges of values of kinetic parameters that support particular model behaviors. For example, oscillations in gene transcription rates might only be produced if the binding of a TF to DNA was appreciable, but not too strong.

The dynamic-systems approach is often preferred over the logical-network approach because it is more accurate. Gene expression rates are continuous variables rather than ideal ON–OFF switches. However, computer integration of dynamic-systems models uses much shorter time steps. Therefore, simulations require much more computer time. Because of this disadvantage, logical-network models may sometimes represent the only practical alternative for modeling large gene networks, or for modeling lengthy processes such as development of tissues or organisms.

Models combining both the logical-network and dynamic-systems approaches have also been developed (McAdams and Shapiro, 1995; Kerszberg and Changeux, 1998; Yuh et al., 1998). With this hybrid approach, biochemical processes that have clear thresholds of activation (e.g., gene activation over a narrow concentration range of an effector molecule) are represented by ON–OFF switches. As an example, if the concentration of a transcription factor is greater than a threshold value, the gene is ON; otherwise, it is OFF. Activation over a broad effector concentration range would be represented instead as a continuous, graded function. Differential equations are included for the rates of change of mRNA and protein concentrations. A particular advantage of the hybrid approach is that clear conceptual distinctions are drawn between activation or repression functions that can be modeled as logical switches, versus activation or repression over a broad range of an effector.

Qualitative Models of Simple Gene Networks

It is of interest to review a few relatively simple, recent models that incorporate typical schemes of gene regulation by TFs. These models use small systems of differential equations to follow the concentration time courses of proteins and mRNAs produced from one or a few genes. Sometimes only autoregulation of a single gene is considered. Studies of such models are expected to provide insights into mechanisms underlying behaviors of more complex gene networks, affected by many TFs and not amenable to intuitive understanding.

Multistability and Time Delays Allow for State- and History-Dependent Responses to Stimuli

Model gene networks often exhibit steady states of fixed gene transcription rates and gene product concentrations. These states can be stable or unstable. "Stable" means that a small movement of the system variables away from the steady state will be followed by relaxation back to that state. "Unstable" means that a small movement of the system variables away from the steady state will lead to further movement away from that state. The term "multistability" refers to the coexistence of multiple stable steady states for a single set of model parameter values. Each state remains stable in the face of disturbances by relatively weak stimuli, but strong stimuli can induce transitions between the coexisting states.

Multistability within a gene network requires a positive-feedback loop (Thomas, 1994; Thomas et al., 1995). In such a loop, a gene product activates, directly or indirectly, transcription of its own gene. A direct loop occurs if a TF activates its own gene. Such TFs include Jun (Angel et al., 1988), myogenic helix-loop-helix proteins (Thayer et al., 1989), and probably CREB, which has CREs in its own promoter region (Meyer et al., 1993; Sassone-Corsi, 1995). An indirect loop occurs if, for example, two TFs repress the transcription of each other's genes. In this case, an increase in the concentration of the first TF will repress the second gene, causing a further increase in transcription of the first gene (see also Figure 5C and below).

Models of gene networks, or of autoregulation of one gene, have been used to investigate which biochemical interactions can support multistability (Keller, 1995; Wolf and Eeckman, 1998). Figure 3A shows a simple example. This one-gene "network" illustrates the effect of a direct positive-feedback loop in which a single transcription factor (TF-A) activates its own transcription. A time delay of 1–2 hr is included for the time between the transcription of *tf-a* and the appearance of TF-A protein. TF-A forms a homodimer. The *tf-a* gene incorporates a responsive element, which we term a TF-RE. Transcription of *tf-a* is increased when TF-A homodimers bind to the TF-RE. TF-A dimerization and resultant cooperativity adds an important nonlinearity to the positive-feedback loop that enables it to support complex system behavior (see discussion below).

The synthesis of TF-A is regulated by stimuli that modify phosphorylation of TF-A because only phosphorylated TF-A dimers are assumed to be effective at activating *tf-a* transcription. First-order degradation of TF-A protein ($-k_d$ [TF-A]) is included to allow the system to reach a steady state. The system of Figure 3A can then be described by a simple first-order differential equation in which the net rate of change of gene product concentration ([TF-A]) is equal to the synthesis rate minus the degradation rate. This single equation can yield relatively rich and complex behavior.

For example, this model exhibits bistability, i.e., the total concentration of TF-A protein ([TF-A]) can be in one of two stable steady states. The existence of two steady states can be explained as follows. In the lower

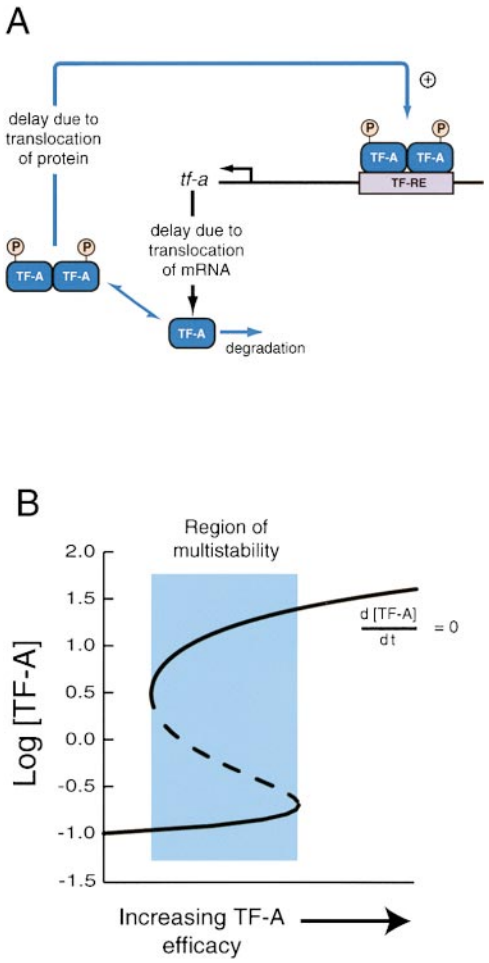

Figure 3. Bistability in a Simple Autocatalytic Model of Gene Regulation

(A) A single transcription factor, TF-A, activates transcription of its own gene when bound to a TF-RE. The sum of time delays between *tf-a* transcription and the appearance of TF-A protein is set to 110 min.

(B) Bistability in the model of (A). Concentration units are micromolars. If the efficacy of TF-A dimers in activating *tf-a* transcription lies within the blue region, two stable steady-state solutions of [TF-A] exist (lower and upper portions of

$$\frac{d[TF-A]}{dt} = 0$$

curve) with an unstable solution between (middle, dashed portion). Outside this region there is a single stable steady-state solution. Modified from Smolen et al. (1999).

state, the low concentration of TF-A protein gives essentially no activation of the positive-feedback loop. A small, basal rate of *tf-a* transcription is balanced by a degradation rate, which is also small since it is proportional to [TF-A]. As [TF-A] is increased above the lower state, the regenerative process of positive feedback occurs, with TF-A dimers activating *tf-a* transcription. Thus, transcription and TF-A protein synthesis can occur at a rate that exceeds TF-A degradation. At a high concentration of TF-A protein, the positive feedback saturates as the TF-RE near the *tf-a* gene becomes fully occupied with TF-A dimers. A maximal transcription rate is approached. The TF-A degradation rate, however, still increases linearly with [TF-A]. Thus, at a high value of [TF-A], the synthesis and degradation rates of TF-A will again come into balance, creating the second steady state.

If the model is in one of the two stable steady states, small perturbations of *tf-a* expression only transiently change [TF-A], and afterwards [TF-A] returns to its original value. Figure 3B illustrates the bistability for a range of values (blue region) of one parameter—the efficacy with which TF-A dimers activate *tf-a* transcription. The efficacy could be modulated by stimulus-induced phosphorylation of TF-A. The black curve traces the values of [TF-A] that remain fixed in time. Within the blue region, the upper and lower solid portions of the curve correspond to the stable steady states. These are separated by an unstable steady state (dashed portion of curve). Since any small disturbance moves the system away from the unstable state, the system does not spend a significant amount of time near it.

One stable state, with a low concentration of TF-A protein ([TF-A]), corresponds to only slow, basal *tf-a* transcription—with very little activation of transcription by TF-A protein. However, suppose a stimulus briefly caused a large increase in *tf-a* expression. This could raise the concentration of TF-A monomers and dimers sufficiently to initiate the positive-feedback loop in which TF-A dimers activate *tf-a* transcription. Counterintuitively, after cessation of the stimulus this positive feedback could indefinitely sustain the system in the upper stable state, with high [TF-A] and strongly activated *tf-a* expression. Figure 4A (upper panel) illustrates switching of this model between the stable states by large perturbations. The model is initially in a steady state of low [TF-A]. A temporary increase in the efficacy of TF-A in activating transcription (left bar) causes a delayed increase of [TF-A] to an upper steady state. The increase does not occur until the time delay of 1–2 hr between the transcription of *tf-a* and the appearance of TF-A has passed. A temporary decrease in the activity of TF-A (right bar) causes a transition back to the lower steady state.

The response of this model to smaller, brief perturbations differs dramatically between states. In the middle panel of Figure 4A, the model is in the lower steady state of [TF-A]. A brief increase in the activity of TF-A (arrow) gives an increase in [TF-A] following the time delay. In the lower panel, the model is in the upper steady state, and the same brief increase in TF-A activation (arrow) causes a much larger increase in [TF-A] (note change in scale of y axis). This difference in response amplitude illustrates how a change in the state of a gene network might "prime" a cell to respond more vigorously to subsequent stimuli.

Time delays in gene networks could strongly affect stimulus responses. With a delay, the response of the positive-feedback loop of Figure 3A to a stimulus is often seen to consist of several "echo" perturbations in [TF-A], with decreasing amplitude (Figure 4). With a delay, the response of a model to a stimulus is history dependent. History dependence means that the response depends on the values of model variables at times previous to the response. Figure 4B illustrates history dependence for the positive-feedback loop of Figure 3A. When the system is in the upper steady state, a brief increase in transcription rate produces a large

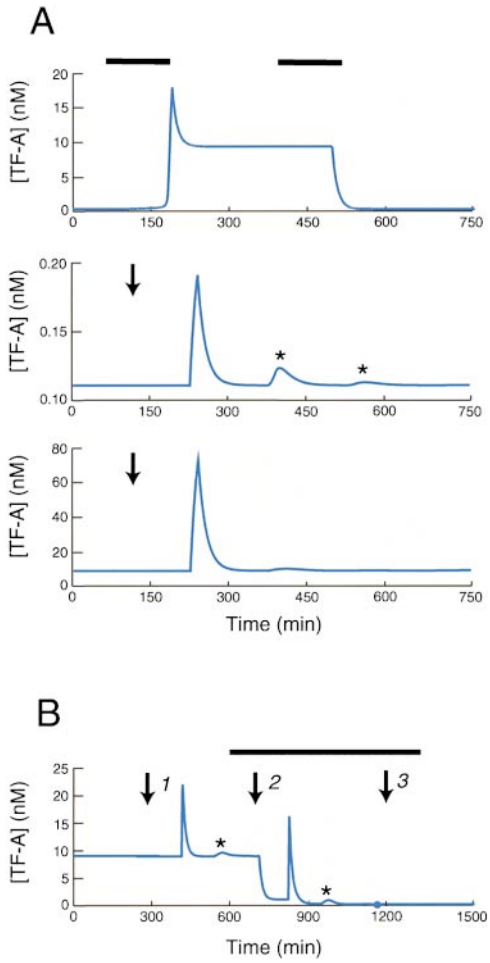

Figure 4. State- and History-Dependent Responses to Stimuli

(A) (Top) The model is initially in a steady state of low [TF-A]. A stimulus from t = 70–200 min is modeled as a temporary increase in the ability of TF-A to increase transcription (left bar). Approximately one time delay (110 min) after the beginning of the stimulus, a transition to an upper steady state results. A second stimulus (right bar)—a decrease in the activity of TF-A (from t = 400–530 min)—causes a delayed transition back to the lower steady state.

(Middle) The model is in the lower steady state. A brief increase in the activity of TF-A at t = 120 min (arrow) gives an increase in [TF-A] following the time delay. Successive smaller increases in [TF-A] labeled by an asterisk are "echo" perturbations.

(Bottom) The model is in the upper steady state, and the same brief increase in TF-A activation (arrow) causes a much larger increase in [TF-A] (note change in scale of y axis).

(B) The model is initially in the upper steady state. At t = 290 min (arrow labeled 1), the degree of TF-A activation is increased for 2 min. Following a time delay (130 min), there is a large increase in [TF-A] at t = 420 min. At t = 600 min the activation of TF-A is decreased to a new baseline (bar). After a delay, at t = 730 min, [TF-A] decreases most of the way to the lower steady state. At t = 700 min (arrow labeled 2), the activation of TF-A is increased for 2 min. A large increase in [TF-A] results at t = 830 min. At t = 1200 min (arrow labeled 3), TF-A activation is increased for 2 min, but the increase in [TF-A] is imperceptible. Small increases in [TF-A] labeled by an asterisk are "echoes."

excursion above this level (following arrow 1). In contrast, when the system is in the lower steady state, the same stimulus produces only a small excursion (response to arrow 3; too small to see). At time t = 600 min, a persistent decrease in the fraction of TF-A able to activate transcription was induced as denoted by the bar. At time t = 730 min, following the time delay, the system began to respond to the persistent change (bar) by shifting to the low steady state of [TF-A]. Meanwhile, at time t = 700 min, a brief stimulus 2 had been applied—the same stimulus as applied at arrows 1 and 3. After a delay, at time t = 830 min, the system responded to stimulus 2 with a large, transient increase in [TF-A]. The response to stimulus 2 illustrates history dependence. Even though the response to stimulus 2 occurred while the system was near the low steady state, the response is large because the system was at the upper steady state at the time of stimulus 2 (contrast with the response to stimulus 3).

Conditions Permitting Oscillations in Gene Network Models

Oscillatory biological phenomena are common and sometimes depend on periodic variation of transcription. Two examples are circadian rhythms (reviewed by Reppert, 1998) and oscillations of CREB mRNA in mammalian secretory cells (Walker et al., 1995). Feedback within gene networks appears to be essential for some instances of oscillations in transcription. Models of circadian rhythmicity often rely on cyclic repression by one or a few genes of their own transcription (e.g., *per* and *tim* in *Drosophila*, or *freq* in *Neurospora*; see below). Oscillations of CREB mRNA probably rely upon positive and negative feedback interactions within the CREB gene family (Figure 1; Walker et al., 1995).

Negative feedback is commonly understood to be important for sustaining oscillations of gene transcription rate. A negative-feedback loop occurs if a gene product directly or indirectly represses transcription of its own gene. A direct loop occurs if a TF represses its own transcription, as occurs with the ICER protein (Molina et al., 1993) and possibly Fos (Sassone-Corsi et al., 1988). In the absence of a time delay, a negative-feedback loop acting alone can support oscillations in transcription rate, but only under restrictive conditions (Thomas et al., 1995). With a time delay, oscillations can more easily be generated by negative feedback alone (Thomas et al., 1995). Generally, systems with only positive feedback do not exhibit oscillations. The reason can be intuitively appreciated by considering the simple model of Figure 3A. Positive feedback can raise the system into the upper steady state, with high [TF-A] and high *tf-a* transcription rate, but there is no opposing, negative-feedback process to bring the concentration of [TF-A] back down. More formally, Smith (1987b) and Smolen et al. (1999) have given mathematical arguments ruling out oscillations for a broad class of systems with only positive feedback, with or without time delays. One condition might allow a positive-feedback system to exhibit oscillations—a limiting rate of supply of a biosynthetic precursor, such as a nucleotide or amino acid. Oscillations might then be generated as the precursor was alternately overused and replenished. In biochemical models of the glycolytic pathway, an analogous mechanism generates oscillations in the rate of metabolite flux through the pathway (Goldbeter, 1995). Here, a key enzyme, phosphofructokinase, is activated by its product, forming a positive-feedback loop. However, a constant limiting rate of substrate supply is assumed.

Gene network models incorporating both negative and positive feedback can readily produce oscillations

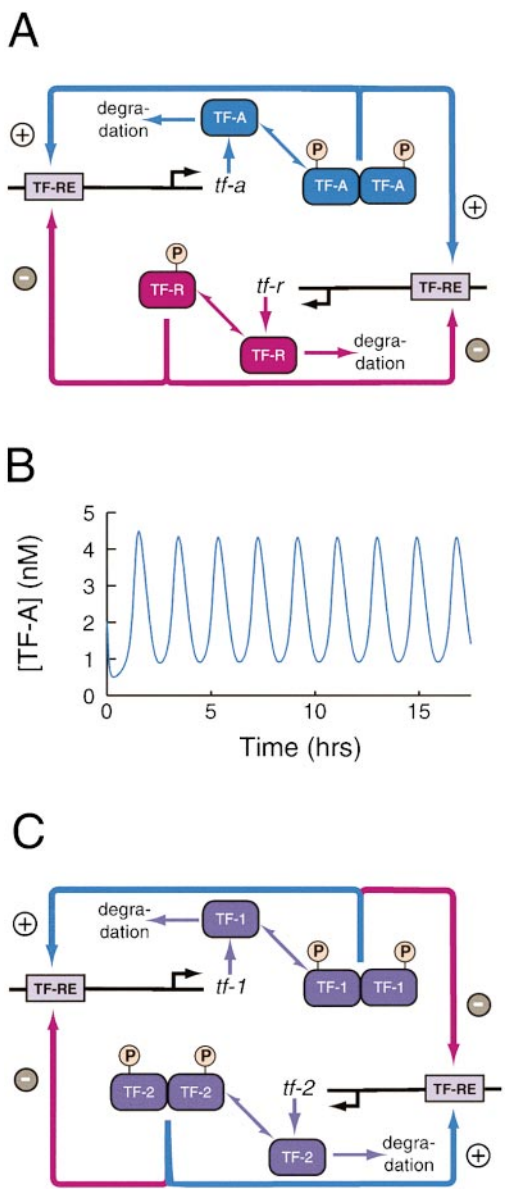

Figure 5. Two Simple Gene Regulatory Schemes that Yield Oscillations and Bistability

(A) Negative feedback yielding oscillations. The transcription factor TF-A activates transcription of its own gene and that for TF-R, whereas TF-R represses transcription of both the TF-A and TF-R genes by competing for DNA binding with TF-A. Only homodimers of TF-A can regulate transcription.

(B) The model of (A) produces oscillations in [TF-A] (and of [TF-R]; not shown) over a wide range of parameters. This model did not incorporate a delay.

(C) Positive feedback that yields bistability. The transcription factor TF-1 activates transcription of its own gene and represses transcription of the gene for TF-2, whereas TF-2 activates its own gene and represses that for TF-1. For both TFs, repression is by competitive binding to DNA, and only homodimers can regulate transcription.

and even more complex behaviors, such as quasiperiodic or chaotic variations in gene transcription rates (Mestl et al., 1996). As a simple example in Figure 5A, a gene for a transcriptional repressor, *tf-r*, was introduced into the previous model of Figure 3A (Smolen et al., 1999). When the TF-A dimer bound to a nearby TF-RE, the rate of *tf-r* transcription was increased. Monomers of TF-R compete with TF-A dimers for binding to TF-RE, thereby repressing transcription of *tf-a* and *tf-r*. Robust oscillations in [TF-A] and [TF-R] were readily generated by this model (Figure 5B).

Oligomerization of TFs Helps to Generate Complex Dynamics

Eukaryotic TFs commonly dimerize or oligomerize before binding to response elements (Kouzarides and Ziff, 1988; Sassone-Corsi, 1995). Early investigators of gene networks recognized that oligomerization of TFs could steepen feedback (increase the Hill coefficient). In general, sharpening feedback in this way favors the occurrence of behaviors associated with nonlinearity, such as multistability and oscillations (Griffith, 1968a, 1968b). Recent models have emphasized this theme. For example, Keller (1995) discussed four multistable models exemplifying typical genetic regulatory schemes that involved homo- or heterodimerization of transcription factors. In these models, multistability was not observed in the absence of TF dimerization. For the models of Figures 3A and 5C, dimerization of transcription factors is also essential for multistability (Smolen et al., 1998). In general, however, TF oligomerization is not an absolute requirement for multistability. Keller (1995) derived conditions for multistability in two schemes involving only TF monomers. In the first scheme, a monomeric TF activates its own transcription by binding to multiple DNA sites. In the second scheme, two monomeric TFs each repress transcription of the other.

Kerszberg (1996) and Kerszberg and Changeux (1994) have developed simple models relying on oligomerization of TFs to illustrate possible mechanisms of spatial organization in developing embryos. In such models, dimers of TFs are used to interpret morphogenetic gradients. These authors have most recently (Kerszberg and Changeux, 1998) modeled aspects of neural tissue development. Within this model, bistability of gene expression rates is created in the context of competition between homodimers of two TFs, each of which activates its own expression while repressing the expression of the other TF. The dimerization and bistability are essential for creating well-defined spatial boundaries of neural versus nonneuronal tissue. Figure 5C illustrates this regulatory scheme, which has also been studied by Keller (1995). Keller (1995) points out that the scheme of Figure 5C constitutes an indirect positive-feedback loop, because an increase in the first TF, by repressing the second, favors a further increase in the first.

Similarly, in models of oscillations in transcription rates within gene networks, oligomerization of TFs is often an essential component. For example, Leloup and Goldbeter (1998) and Goldbeter (1995) constructed models of circadian rhythmicity that rely on dimerization of TFs, as well as negative feedback. Also, in the model of Figure 5A with activation and repression of transcription, dimerization of TF-A adds a nonlinearity that is essential for oscillatory behavior (Smolen et al., 1998). Dimerization steepens the positive feedback of TF-A upon transcription and enhances its regenerative nature, thus giving the sharp upstrokes of TF-A and TF-R protein that form the first phase of each oscillation. Only a steady state is seen if dimerization is neglected. Together, the above results emphasize how oligomerization of TFs promotes complex transcriptional dynamics.

Competition between Transcriptional Activators and Repressors Could Yield Optimal Stimulus Frequencies for Transcription

Experimental systems have provided evidence that the rate of transcription can have a complex dependence on stimulus frequency. For example, in cultured neurons, c-*fos* transcription was nearly tripled by bursts of six electrical stimuli (10 Hz) with an interburst interval of 1 min, but transcription was not significantly affected by bursts of 12 stimuli with an interburst interval of 2 min (Sheng et al., 1993). Also, transcription of the cell adhesion molecule L1 was strongly repressed by continuous 0.1 Hz electrical stimulation, but not significantly affected by 0.3 Hz stimulation (Itoh et al., 1995).

Frequency-dependent transcription might also help to explain aspects of LTM, the formation of which has been shown to be frequency dependent in some cases. An example comes from an experimental model of olfactory learning in *Drosophila* (Tully et al., 1994). In this system, repeated training trials that had a relatively long intertrial interval (ITI) yielded more LTM than equivalent training trials presented using a short ITI, even given the same total training time. Optimal stimulus frequencies may also exist for some examples of task learning by humans (Kientzle, 1946).

Recently, modeling has been used to describe possible mechanisms for generating frequency-dependent transcription. A kinetic scheme with activation and repression of transcription (Figure 5A) was extended to include more detailed descriptions of TF phosphorylation and dephosphorylation (Smolen et al., 1998). By adjusting the kinetics of phosphorylation and dephosphorylation, the activity of the TFs in response to stimuli was made such that a window of stimulus frequencies activated TF-A but hardly activated TF-R. An "output" gene was transcribed only when TF-A bound to its TF-RE. Dimers of activated TF-A and TF-R competed for binding to this site. The model predicted a stimulus frequency at which transcription of the output gene was maximal. In a similar model, an optimal stimulus frequency was created by competition between dual phosphorylations of a single TF. Here, the first site was readily phosphorylated by low-frequency stimuli and allowed the TF to activate transcription, while the second site was phosphorylated only by high-frequency stimuli and countered the effect of the first. These models illustrate a qualitative principle. Two competing processes whose kinetics endow them with different sensitivities to stimuli, such as activator and repressor phosphorylation, can compete to maximize or minimize transcription at specific stimulus frequencies.

Modeling Macromolecular Transport with a Time Delay, or Diffusively, Can Profoundly Affect the Dynamics of Gene Networks

Several studies have compared the behavior of model gene networks that use different methods for describing intracellular macromolecular transport (Busenberg and Mahaffy, 1985; Mahaffy et al., 1992; Smolen et al., 1999). The methods used most often are time delays and diffusion. In models that use diffusion, slowing diffusion is usually found to dampen oscillations in transcription rates, giving a steady state. By contrast, in models that use time delays, increasing the time delays (which is analogous to slowing diffusion) tends to destabilize steady states and promote oscillations.

To examine whether the dynamics of the simple models of Figures 3A and 5A would be altered when transport was modeled by diffusion instead of time delays, Smolen et al. (1999) modeled diffusion of *tf-a* and *tf-r* mRNA and protein in a spherically symmetric cell divided into concentric shells. The positive-feedback loop of Figure 3A still exhibited bistability in [TF-A]. However, in response to a brief stimulus, the time courses of mRNA and protein only showed a single response. They never exhibited the "echoes," or repetitive responses, described in Figure 4A. Diffusion acts to "spread out" the spatial peaks of macromolecular concentration that are necessary for "echoes". By contrast, a time delay preserves peaks because it corresponds to transporting concentration time courses without any change in shape. In the model of Figure 5A, slowing diffusion tended only to suppress oscillations in macromolecular concentrations. In contrast, as the studies cited above would predict, increasing a time delay often created oscillations.

Random Fluctuations in Macromolecule Numbers Can Yield Significant Variability in the Behavior of Genetic Systems

In many genetic systems, the average number of copies of important macromolecules is low (tens to hundreds). In such conditions, randomness in the timing of synthesis or degradation of individual molecules is expected to yield significant fluctuations in copy numbers (Keizer, 1987; McAdams and Arkin, 1999). Transcription of individual eukaryotic mRNAs is in fact known to proceed in a random, irregular manner, which would lead to fluctuations in mRNA concentrations (Ross et al., 1994; Zlokarnik et al., 1998). Random fluctuations in the copy numbers of transcription factors in the nucleus, and in the expression of genes affected by these TFs, might have phenotypic consequences. As a possible example, attempts to reduce morphologic variability in inbred mice have generally found a lower limit. Irreducible variability in this and other species may be due to randomness in the timing of expression events occurring early in embryogenesis (Gartner, 1990).

A reliable Monte Carlo algorithm for simulating fluctuating biochemical reactions has been presented by Gillespie (1977). This algorithm was recently used by McAdams and Arkin (1997) in qualitative models of prokaryotic gene networks with TF dimerization and with regulatory feedback. Figure 6A illustrates an example in which dimers of protein P_1 repress the transcription of a gene p_2 and activate the transcription of another gene, p_3. The system has a single negative-feedback loop in which P_1 dimers repress p_1 transcription. The time taken for P_1 to accumulate was subject to large stochastic variability (Figure 6B). For accumulation of 25 P_1 dimers, the time required for each of 100 simulations varied from $<$10 min to $>$45 min. Consider a situation in which P_1 regulates genes that govern phenotype and P_1 accumulation is triggered by an environmental stimulus of limited duration. McAdams and Arkin (1997) point out that, under the influence of randomness observed in this model, only a subpopulation of cells would accumulate sufficient P_1 during the stimulus to reach a threshold sufficient to effect a change in phenotype. The importance of such variability in many eukaryotic systems remains to be assessed. For example, the extent of neuronal plasticity induced by a given stimulus could be quite variable at the level of a single cell or synapse, but the average of stochastic variability over many neurons—a relevant quantity for assessing the

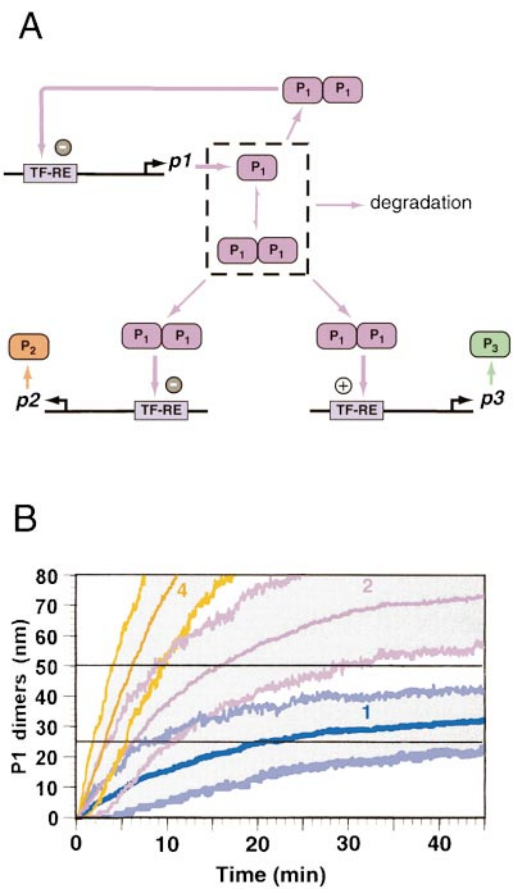

Figure 6. Stochastic Fluctuations in Molecule Numbers Affect the Dynamics of a Simple Gene Network

(A) The kinetic scheme used for simulations of McAdams and Arkin (1997). The product of gene p_1 dimerizes. Both monomeric and dimeric P_1 protein are subject to first-order degradation (dashed box). P_1 homodimer represses its own transcription and that of gene p_2 and activates transcription of gene p_3. All processes of macromolecular synthesis and degradation, as well as binding and unbinding between protein molecules and responsive elements, are stochastic in nature.

(B) Results for 100 stochastic simulations of the onset of P_1 dimer production assuming the kinetic scheme of (A), at p_1 gene dosages of 1, 2, and 4. Plots represent mean $\pm$ 1 standard deviation (gray regions around mean time courses). P_1 concentration has units of nanomolars. Great variability in the amount of time required for individual simulations to reach particular P_1 dimer levels of 25 or 50 molecules (horizontal lines) is evident. Reproduced with permission from McAdams and Arkin (1997).

reliability of memory formation—would be considerably less.

A general physical principle is that the lower the average number of molecules, the greater the relative size and importance of random fluctuations (Keizer, 1987). Therefore, in a multistable genetic system, random fluctuations in molecule numbers would tend to preferentially "mask" the existence of steady states with low average copy numbers of macromolecules. Recently, fluctuations were added to the bistable positive-feedback loop of Figure 3A to investigate the likelihood of such "masking" (Smolen et al., 1999). With a physiologically plausible set of parameter values, the steady state with a lower concentration of TF-A was masked. If a simulation was begun with the model at the lower steady state, it remained nearby for a few hours before fluctuations caused a transition to the upper state. In actual genetic systems, the strength of "masking" is strongly parameter dependent, because biochemical rate constants determine fluctuation characteristics (Hasty et al., 2000). For example, in an engineered bistable genetic switch in *E. coli*, the lower state was found to be more stable, even though fluctuations were appreciable and caused cell-to-cell variability in the location of the switch thresholds (Gardner et al., 2000). This special result can only be found if the lower state (with fewer macromolecules on average) is more strongly attracting than the upper state. That is, the upper state must be only weakly stable, so that a relatively small perturbation suffices to move the system from the upper to the lower state.

Future studies that model specific gene networks should assess the likely importance of stochasticity, including the effect of stochasticity on the stabilities of steady states. However, deterministic models will remain important because data are often insufficient to model stochastic fluctuations in detail. Data used to construct deterministic models are often obtained from large and reproducible responses to stimuli. Such models may, therefore, be expected to predict reliably the responses to new stimuli of comparable magnitude.

In summary, mathematical models of simple gene networks, such as those considered above, have revealed associations between common biochemical elements and dynamic behaviors (Table 1). Specific features of the biochemical pathways, such as feedback loops, dimerization, and time delays, support complex dynamics, such as bistability and oscillations. These relationships may give insight into the behavior of newly characterized gene networks.

Models of Specific Gene Networks

Models of particular gene networks need to be based, insofar as possible, on measured values of biochemical parameters. Use of parameter values that do not describe the in vivo situation can lead both to erroneous predictions of quantities, such as the size of stimulus responses, and to erroneous predictions of qualitative dynamic behaviors, such as oscillations or multistability. However, for these complicated biochemical systems, obtaining good estimates of all important parameters proves difficult in practice. For example, time courses of gene expression rates can be fitted with kinetic models, but the set of parameter values thereby obtained may not be unique or strongly constrained.

Because the somewhat less complex structure of prokaryotic cells tends to simplify the acquisition of data to constrain model parameters, prokaryotic gene network models have been more common than eukaryotic models. The *E. coli*–λ phage system is perhaps the best characterized prokaryotic gene network. Two phage proteins competitively control a switch point that determines whether an infected *E. coli* cell follows a lysogenic or lytic pathway. Recently, Arkin et al. (1998) constructed a detailed model of this genetic switch. This model included random timing of individual macromolecular synthesis and degradation events, and successfully simulated experimental data for the fractions of cells undergoing lysogeny versus lysis under different environmental conditions. Other types of prokaryotic genetic switches include those based on competition

Table 1. Associations in Gene Network Models between Common Biochemical Elements and Dynamic Behaviors

Biochemical Element	Favored Behavior
positive feedback	multistability
	oscillations possible only if supply of biosynthetic precursors is rate limiting
	state-dependent responses ("priming")
negative feedback	oscillations
biochemical time delays	complex responses to stimuli
	oscillations
	"echoes"
competition between gene activation and repression	optimal stimulus frequencies for extrema in transcription rates
protein oligomerization	multistability
	oscillations
	optimal stimulus frequencies
stochastic fluctuations in molecule numbers	destabilization of some steady states
	randomness in responses to stimuli
	random outcomes of processes dependent on genetic regulation

among transcription factors for binding to promoters (Hicks and Grossman, 1996) and on inversion of DNA segments (Dorman, 1995). Stochastic variations in the operation of such switches are expected to contribute significantly to phenotypic variability (Laurent and Kellershohn, 1999).

Switch-like responses to stimuli are also common in eukaryotic systems. For example, the MAP kinase pathway, which affects the expression of a variety of genes including those of the CREB family of TFs, can exhibit a very steep stimulus threshold for elicitation of a response, so that it appears to respond in an "all-or-nothing" fashion (Huang and Ferrell, 1996). Stochastic variability in operation of such a switch could affect, for example, whether a stimulus elicits gene transcription and neurite growth in a particular neuron. The degree of variability expected depends strongly on rate constants and other parameters. Thus, its assessment requires experimental determination of kinetic parameters in combination with quantitative modeling.

The development of models of eukaryotic systems has been hindered by the difficulty of obtaining biochemical data. One eukaryotic network that has been modeled in some detail is the regulation by TFs of the gene *endo16*, necessary for development of sea urchin embryos (Yuh et al., 1998). In this network, at least 15 TFs regulate *endo16* transcription, and there are 8 distinct TF-REs upstream from the basal promoter. To obtain data on the time course of expression of the various regulatory regions, Yuh et al. (1998) constructed TF-REs attached to reporter genes. Time courses for the expression of each construct were determined during development. The results of these experiments helped to clarify the function of each TF-RE. For three TF-REs (those for which the most extensive kinetic data were available), their degrees of occupancy by TFs were modeled as real, continuous variables. The rate of *endo16* transcription was also considered a continuous variable. The other five TF-REs had their occupancies modeled as simple switches—either OCCUPIED (1) or VACANT (0). The model sufficed to qualitatively illustrate how the *endo16* regulatory elements integrate environmental information via occupancies of TF-REs to produce a well-defined time course of the rate of *endo16* transcription. However, it is evident that much remains to be done in understanding this system. The roles of most of the regulatory TFs have yet to be determined, and relationships between TF concentrations and their effects are unknown.

Recently, Kerszberg and Changeux (1998) have modeled aspects of neural tissue development in vertebrate and insect embryos. Within this model, bistability of gene expression rates is created by competition between homodimers of two TFs, each of which activates its own expression while repressing the expression of the other TF (Figure 5C). These TFs are qualitatively analogous to known TFs encoded by the *enhancer-of-split* and *achaete-scute* gene families in *Drosophila*. A mechanism for intercellular signaling, qualitatively similar to the Notch–Delta receptor–ligand system, is assumed to modulate expression of the "enhancer-of-split" TF. Well-defined spatial boundaries of neural versus epithelial tissue result, with cells on each side of the boundary in a different gene expression state. These gene expression states are assumed to be epigenetically inherited by daughter cells, preserving established tissue boundaries during further development. The model also includes other TFs known to regulate neurogenesis, and the authors propose that cell adhesion and motion are regulated by genes critical for bistability and switch-like behavior (e.g., *Notch* or *achaete-scute*).

Another class of eukaryotic models has been developed to help understand the transcriptional regulation responsible for circadian rhythms. Circadian rhythms are found in all organisms adapted to day–night cycles, whether they be cyanobacteria, fungi, or mammals. In each species, a few genes appear to act as essential core genes underlying expression of the rhythm (Reppert, 1998). In a negative-feedback loop, one or more of the core gene products represses its own transcription after a delay of a few hours (Dunlap et al., 1999). Several hours later, these gene products are degraded, beginning the next cycle of the oscillation. Identified core genes that repress their own transcription include *per* and *tim* in *Drosophila* (for a recent review, see Young, 1998) and *freq* in the fungus *Neurospora* (Merrow et al., 1997). Much modeling of circadian oscillations has been

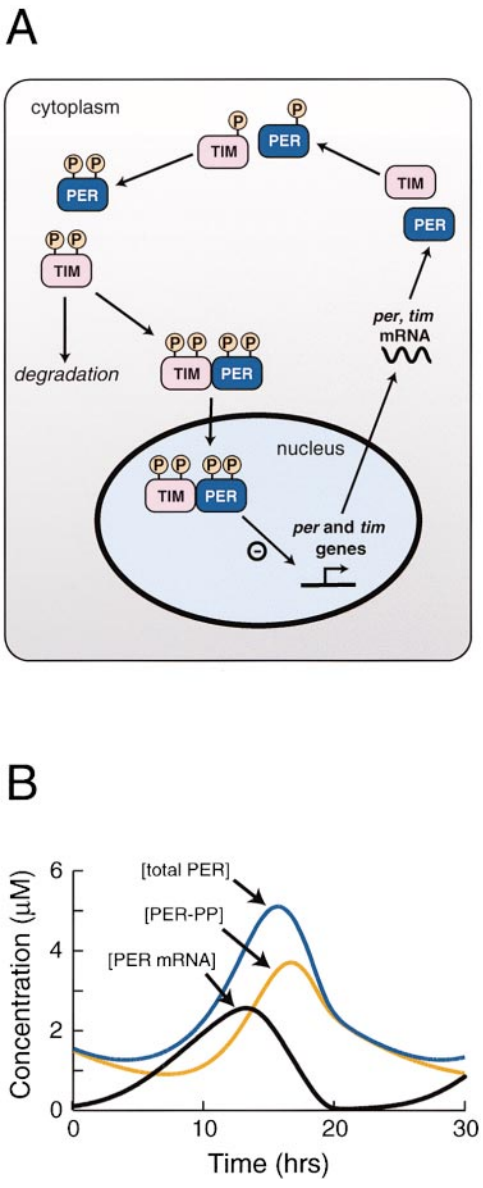

Figure 7. Simulation of Circadian Oscillations in *Drosophila* by the Model of Leloup and Goldbeter (1998) Incorporating Known Negative-Feedback Interactions

(A) The *Drosophila* PER and TIM proteins form a heterodimer, which can translocate to the nucleus. Translocation leads to repression of *per* and *tim* transcription. The PER and TIM proteins are phosphorylated over a period of hours and are then degraded. In this model, two phosphorylations of PER and TIM are assumed necessary for degradation.

(B) Simulation of circadian oscillations. Only time courses of *per* mRNA and PER protein are shown (time courses of *tim* mRNA and TIM protein are similar). [PER-PP] denotes the concentration of all species (free and in heterodimers) of doubly phosphorylated PER.

qualitative and not specific to any organism (e.g., Scheper et al., 1999). However, specific models of the *Drosophila* oscillator have been developed (Goldbeter, 1995; Leloup and Goldbeter, 1998) that successfully reproduce circadian oscillations in gene expression as well as the response of the oscillator to brief light stimuli or light–dark cycles. Figure 7A illustrates key features of the model of Leloup and Goldbeter (1998). PER and TIM heterodimerize, and the heterodimer represses *per* and *tim* transcription. After multiple slow phosphorylations, PER and TIM are degraded. In this model, only two phosphorylations are included to avoid excessive complexity. Figure 7B illustrates circadian oscillations simulated by this model. In concordance with experimental data (Lee et al., 1998), the peak of *per* mRNA concentration leads the peak of total PER protein by about 4 hr, and the time course of total PER leads that of doubly phosphorylated PER by a few hours. This model is also able to reproduce *Drosophila* phase–response behavior by predicting the shift of circadian rhythm caused by brief light pulses delivered at different phases of the circadian cycle.

However, recent advances in experimental knowledge of circadian genes in *Drosophila*, *Neurospora*, and other species have made it evident that current models are missing important elements. For example, transcriptional autorepression appears to be indirect. Core gene products appear to bind and sequester other proteins necessary for core gene expression. In *Drosophila*, for example, *clock* and *cycle* are essential for rhythmicity, and their products form a heterodimer that activates *per* transcription and is bound by the PER-TIM heterodimer (Lee et al., 1998). In addition, transcription of *clock* is activated by PER and/or TIM, and the *clock* gene product represses *clock* transcription (Glossop et al., 1999; Bae et al., 2000). These regulatory complexities have not been modeled. Experimental knowledge of the *Neurospora* oscillator is also progressing rapidly (Merrow et al., 1997; Luo et al., 1998). Here, the proteins WC-1 and/or WC-2 may activate *freq* transcription and may be sequestered by FREQ (Dunlap, 1998). Thus, the time may be ripe for constructing a detailed model of the *Neurospora* oscillator. Preliminary efforts have been reported (Roenneberg and Merrow, 1998; Smolen et al., 1999, Soc. Neurosci., abstract).

More generally, it has been suggested that positive as well as negative feedback is essential for the operation of circadian oscillators (Crosthwaite et al., 1997). Although oscillations can be sustained by negative feedback alone, the presence of positive feedback might increase the robustness of oscillations to parameter variation and might modulate responses to stimuli. In systems with multiple feedback interactions, the relative importance of each can only be assessed by modeling. Preliminary efforts (Smolen et al., 1999, Soc. Neurosci., abstract) are underway to examine the effect of adding a known positive-feedback interaction—activation by PER-TIM heterodimers of *clock* transcription—to the model of Leloup and Goldbeter (1998) for the *Drosophila* oscillator. Finally, the period and amplitude of circadian oscillations should be rather robust to stochastic fluctuations in the copy numbers of important macromolecules. This constraint is being used to assess the realism of proposed molecular models of circadian oscillators (Barkai and Leibler, 2000).

Specific Issues that Require Further Investigation

Modeling and Experimental Approaches Can Be Employed Concurrently to Understand Gene Networks Important for Learning and Memory

Figure 1 illustrated a particular gene network based on CREB and related transcription factors that appears to be important for the formation of LTM (Dash et al., 1990;

Bartsch et al., 1995, 1998; Dash and Moore, 1996; Finkbeiner et al., 1997; Guzowski and McGaugh, 1997; Lamprecht et al., 1997; Glazewski et al., 1999). Aspects of this gene network resemble Figure 5A. A negative-feedback loop exists, in which ICER is induced by CREB and then binds to CREs and represses transcription of CREB. A positive-feedback loop also appears to exist, in which CREB binds to CREs and activates its own transcription.

Experimental data concerning the kinetics of activation or expression of these TFs have not yet been integrated into a detailed model. Such a model might, for example, help to predict the efficacy of the formation of LTM during different training protocols. A molecular mechanism for creating an optimal frequency for the formation of LTM was proposed by Yin et al. (1995). These authors hypothesized that two critical TFs, an activator (TF-A) and a repressor (TF-R), were activated by stimuli, perhaps via phosphorylation. TF-A was then able to activate transcription of key "target" genes necessary for the formation of LTM, whereas TF-R competed with TF-A to repress the target genes. TF-R is dephosphorylated more rapidly than TF-A. Thus, during low-frequency repetitive stimuli, the average degree of TF-R phosphorylation was less than that of TF-A. Low-frequency stimuli would therefore tend only to activate TF-A and activate transcription of "target" genes necessary for the formation of LTM. With higher frequency stimuli, TF-R would also activate and stay active because there is insufficient time for its dephosphorylation between stimuli. Thus, TF-R would antagonize TF-A and repress "target" gene transcription. To verify the plausibility of this hypothesis, Smolen et al. (1998) modified the model of Figure 5A to include time-dependent kinetics of TF-A and TF-R phosphorylation. It was found that kinetic parameters could readily be chosen to allow sharply peaked, optimal stimulus frequencies for maximal transcription.

The CREB family of proteins contains both activators (such as CREB) and repressors (such as CREB2) whose activity is modulated by phosphorylation and that could implement the hypothesis of Tully et al. (1994). However, with respect to the repressor CREB2, recent evidence obtained in *Aplysia* suggests that phosphorylation *reduces* its repressing activity (Martin et al., 1997). But, an optimal stimulus frequency for transcription could also be generated in this gene network from other competing processes. For example, CREB has both activating and inhibiting phosphorylation sites (Ser-133 and Ser-142) (Sassone-Corsi, 1995). If these sites were phosphorylated by stimuli of different frequencies, frequency tuning of CREB activity could emerge.

The Frequency and Functional Consequences of Gene Network Behaviors Such as Multistability Have Not Yet Been Fully Assessed

Many architectures for gene networks have been observed; those discussed in this review are merely a sample. A few examples of network features not discussed include: heterodimerization of transcription factors in two separate pathways with a shared third TF (Hunter et al., 1996), convergence of signaling pathways through distinct TFs onto a single gene (Howard and Maurer, 1995), and conditional regulation by a single TF (e.g., enhancement of constitutive transcription of the dopamine β-hydroxylase gene by the YY1 TF, which in the same preparation suppresses cAMP-induced transcription; Seo et al., 1996). Also, TF-induced DNA bending can influence the actions of other TFs, and interactions between TFs are frequently promoter specific, depending on the relative positioning of DNA binding sites (Fry and Farnham, 1999). Mechanisms of genetic control encompass not only transcriptional regulation by TFs but also regulation of translational and posttranslational processes. This diversity of motifs for transcriptional regulation suggests that behaviors dependent on nonlinearity and on feedback interactions (such as multistability, oscillations, and frequency tuning) may be important in many gene networks.

In considering this suggestion, several authors have addressed possible functional consequences of multistability in particular. Brief exposure of cells to a neurotransmitter or hormone, or brief periods of electrical activity and calcium influx in neurons or other excitable cells, might "trigger the switch" and lead to long-lasting changes in gene expression. Because of biochemical time delays, the occurrence of state transitions is likely to be very dependent on stimulus duration as well as amplitude. In neuronal circuits, induced steady states of gene expression could help foster the synthesis of proteins required for synaptic growth and the formation of LTM. Steady states of gene expression levels could also constitute an important form of mitotically inheritable epigenetic information (Keller, 1994). Furthermore, Thomas et al. (1995) hypothesized that cell differentiation is largely a biological expression of multistability of protein concentrations. Because multistability is a sign of a positive-feedback loop, these authors suggested that genes involved in positive-feedback loops should be examined carefully for potential roles in differentiation. Laurent and Kellershohn (1999) and Keller (1994) have also proposed that switching between multiple stable states of gene expression is an essential component of differentiation.

Modeling Specific Gene Networks Will Become More Feasible with New Methods for Gathering Detailed Data

A variety of new biochemical methods are providing data useful for modeling gene networks. Transfection of cells with reporter genes whose promoters have been chosen to be regulated by specific TFs permits measurements of time courses of transcription (Castano et al., 1996). Analysis of these time courses can yield estimates of kinetic parameters governing the onset and termination of transcription and mRNA degradation.

Another method involves the use of DNA chips. These are high-density arrays of oligonucleotides, with each oligonucleotide chosen to hybridize to a distinct mRNA sequence (Winzeler et al., 1998; Gerhold et al., 1999). This method allows the simultaneous measurement of the time courses of expression of up to ~10,000 genes. These data will prove invaluable for studies of responses to environmental agents, stimuli that induce LTM, or studies of development (e.g., White et al., 1999). PCR amplification and quantitation of mRNAs from tissue samples can also be used to calculate expression time courses of large numbers of genes (Wen et al., 1998). Genes can be grouped into clusters with similar expression profiles (Eisen et al., 1998). Similarity in expression profile suggests similarity in regulation and possibly in function. Thus, potential relations between the functions of new genes and characterized genes can be identified. As an example, Wen et al. (1998) used cluster analysis to characterize spinal cord development. One hundred twelve genes were classified into five temporal episodes

of expression. Cluster analysis revealed several associations between the expression of genes with known and unknown functions. Also, mathematical analysis of the pooled gene expression data revealed pronounced asymmetries, which emphasized that gene expression during development is overall a strongly constrained and ordered process.

Methods for using sets of expression time courses to rapidly construct logical-network models of genetic systems are being developed (Somogyi and Sniegoski, 1997; Wen et al., 1998). In addition, expression time courses might reveal long-lasting state transitions in response to brief stimuli. Such state transitions would suggest multistability and the existence of a positive-feedback loop (Thomas et al., 1995).

There are important limitations to the use of DNA chips or PCR for cluster analysis. These techniques monitor populations of cells and are generally unable to resolve expression dynamics in single cells. If the mRNA is from cells of different types (e.g., if obtained from a tissue sample containing both glia and neurons), the analysis is complicated further. However, some of these limitations can be overcome by using marker genes known to be specific to one cell type. By using cluster analysis, genes associated to the marker genes by similar time courses of expression can be identified. In such a scenario, it is reasonable to assume that the associated genes are expressed in the same cell type as the marker genes and to guess, as a pointer toward further study, that the genes with similar time courses share common regulatory inputs. PCR methods for analyzing gene expression in single cells have begun to be developed, but it appears that at this stage the intrinsic variability of these methods is too high to allow for reliable quantitation of differences in gene expression between individual cells, unless these differences are quite large (>5-fold) (Brail et al., 1999).

We note that within the emerging discipline of proteomics, automated methods are being developed to characterize the amounts and interactions of proteins expressed under a variety of conditions (Ryu and Nam, 2000). These data are a necessary complement to mRNA quantitation because the correlation between mRNA and protein levels is often poor. A major challenge is to refine mathematical techniques to identify functional relationships within large groups of proteins (Marcotte et al., 1999). In constructing models of gene networks, it is desirable to incorporate any data concerning protein time courses. This allows more accurate modeling of transcriptional regulation, and better characterizes network "output"—i.e., overall protein production.

Other data important for constructing accurate models of gene networks are rates of intracellular transport of important macromolecules and information about whether such transport is active or passive. A detailed picture of intracellular transport of mRNA and protein is, in many respects, not yet established. Active transport of mRNAs and proteins has now been demonstrated in numerous cases, often in neuronal processes (Yisraeli et al., 1990; Sabry et al., 1995; King, 1996; Kiebler and DesGroseillers, 2000). However, there is no consensus concerning the extent to which a "typical" mRNA is actively transported. A technique that may facilitate modeling is visualization of macromolecules using fluorescent in situ hybridization (FISH; Femino et al., 1998). This technique could help determine whether specific mRNAs are actively transported or move by diffusion. For example, Femino et al. (1998) monitored an F-actin gene in cultured fibroblasts and observed that mRNAs often departed the gene along definite tracks. This observation suggested an active-transport mechanism in which mRNAs were directed along cytoskeletal elements by motor proteins. But in at least 50% of the cases, the mRNAs appeared to simply diffuse away from the gene.

Further development of experimental techniques is expected to facilitate construction of detailed models of gene networks important for controlling key biological events such as differentiation or the formation of LTM. The responses of tissues and organisms to normal and unusual stimuli will be more thoroughly understood through modeling and experiments that characterize the regulation of relevant genes. Analysis and prediction of responses of organisms to environmental contaminants, or to novel pharmaceutical agents, will be significantly enhanced by collaborative modeling and experiment.

Acknowledgments

We thank A. Angers, J. Chin, L. Cleary, S. Candy, E. Robson, and K. Scholz for their comments on the manuscript. This work was supported by National Institutes of Health grants T32 NS07373, R01 RR11626, and P01 NS38310.

References

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of *Drosophila melanogaster*. Science *287*, 2185–2195.

Angel, P., Hattori, K., Smeal, T., and Karin, M. (1988). The *jun* proto-oncogene is positively autoregulated by its product, Jun−AP-1. Cell *55*, 875–885.

Arkin, A., Ross, J, and McAdams, H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage λ − infected *E. coli* cells. Genetics *149*, 1633–1648.

Bae, K., Lee, C., Hardin, P.E., and Edery, I. (2000). dCLOCK is present in limiting amounts and likely mediates daily interactions between the dCLOCK-CYC transcription factor and the PER-TIM complex. J. Neurosci. *20*, 1746–1753.

Bagley, R.J., and Glass, L. (1996). Counting and classifying attractors in high dimensional dynamical systems. J. Theor. Biol. *183*, 269–284.

Barkai, N., and Leibler, S. (2000). Circadian clocks limited by noise. Nature *403*, 267–268.

Bartsch, D., Casadio, A., Karl, K., Serodio, P., and Kandel, E.R. (1998). *CREB1* encodes a nuclear activator, a repressor, and a cytoplasmic modulator that form a regulatory unit critical for long-term facilitation. Cell *95*, 211–223.

Bartsch, D., Ghiradi, M., Skehel, P., Karl, K., Herder, S., Chen, M., Bailey, C., and Kandel, E.R. (1995). *Aplysia* CREB2 represses long-term facilitation: relief of repression converts a transient facilitation into long-term functional and structural change. Cell *83*, 979–992.

Brail, L.H., Jang, A., Billia, F., Iscove, N., Klamut, H., and Hill, P. (1999). Gene expression in individual cells: analysis using global single cell reverse transcription polymerase chain reaction (GSC RT-PCR). Mutat. Res. *406*, 45–54.

Busenberg, S., and Mahaffy, J. (1985). Interaction of spatial diffusion and delays in models of genetic control by repression. J. Math. Biol. *22*, 313–333.

The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science *282*, 2012–2018.

Castano, J., Kineman, R., and Frawley, L.S. (1996). Dynamic monitoring and quantification of gene expression in single, living cells: a

molecular basis for secretory cell heterogeneity. Mol. Endocrinol. *10*, 599–605.

Collins, F.S., Patrinos, A., Jordan, E., Chakravati, A., Gesteland, R., and Walters, L. (1998). New goals for the U.S. human genome project: 1998–2003. Science *282*, 682–689.

Crosthwaite, S.K., Dunlap, J.C., and Loros, J.J. (1997). *Neurospora wc-1* and *wc-2*: transcription, photoresponses, and the origins of circadian rhythmicity. Science *276*, 763–769.

Dash, P.K., and Moore, A.N. (1996). Characterization and phosphorylation of CREB-like proteins in *Aplysia* central nervous system. Brain Res. Mol. Brain Res. *39*, 43–51.

Dash, P.K., Hochner, B., and Kandel, E.R. (1990). Injection of the cAMP-responsive element into the nucleus of *Aplysia* sensory neurons blocks long-term facilitation. Nature *345*, 718–721.

Dorman, C.J. (1995). DNA topology and the global control of bacterial gene expression: implications for the regulation of virulence gene expression. Microbiology *141*, 1271–1280.

Dunlap, J.C. (1998). Common threads in eukaryotic circadian systems. Curr. Opin. Genet. Dev. *8*, 400–406.

Dunlap, J.C., Loros, J.J., Liu, Y., and Crosthwaite, S.K. (1999). Eukaryotic circadian systems: cycles in common. Genes Cells *4*, 1–10.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA *95*, 14863–14868.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts *in situ*. Science *280*, 585–590.

Finkbeiner, S., and Greenberg, M.E. (1998). Ca^{2+} channel-regulated neuronal gene expression. J. Neurobiol. *37*, 171–189.

Finkbeiner, S., Tavazoie, S.F., Maloratsky, A., Jacobs, K.M., Harris, K.M., and Greenberg, M.E. (1997). CREB: a major mediator of neuronal neurotrophin responses. Neuron *19*, 1031–1047.

Fry, C.J., and Farnham, P.J. (1999). Context-dependent transcriptional regulation. J. Biol. Chem. *274*, 29583–29586.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. Nature *403*, 339–342.

Gartner, K. (1990). A third component causing variability beside environment and genotype. A reason for the limited success of a 30-year long effort to standardize laboratory animals? Lab. Anim. *24*, 71–77.

Gerhold, D., Rushmore, T., and Caskey, C.T. (1999). DNA chips: promising toys have become powerful tools. Trends Biochem. Sci. *24*, 168–173.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. *61*, 2340–2361.

Glass, L., and Kauffman, S.A. (1973). The logical analysis of continuous, non-linear biochemical control networks. J. Theor. Biol. *39*, 103–129.

Glazewski, S., Barth, A.L., Wallace, H., McKenna, M., Silva, A., and Fox, K. (1999). Impaired experience-dependent plasticity in barrel cortex of mice lacking the α and δ isoforms of CREB. Cereb. Cortex *9*, 249–256.

Glossop, N., Lyons, L.C., and Hardin, P.E. (1999). Interlocked feedback loops within the *Drosophila* circadian oscillator. Science *286*, 766–768.

Goldbeter, A. (1995). A model for circadian oscillations in the *Drosophila* Period (PER) protein. Proc. R. Soc. Lond. B Biol. Sci. *261*, 319–324.

Griffith, J.S. (1968a). Mathematics of cellular control processes. I. Negative feedback to one gene. J. Theor. Biol. *20*, 202–208.

Griffith, J.S. (1968b). Mathematics of cellular control processes. II. Positive feedback to one gene. J. Theor. Biol. *20*, 209–216.

Guckenheimer, J., and Holmes, P. (1983). Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (New York: Springer-Verlag).

Guzowski, J., and McGaugh, J.L. (1997). Antisense oligodeoxynucleotide-mediated disruption of hippocampal cAMP response element binding protein levels impairs consolidation of memory for water maze training. Proc. Natl. Acad. Sci. USA *94*, 2693–2698.

Hasty, J., Pradines, J., Dolnik, M., and Collins, J.J. (2000). Noise-based switches and amplifiers for gene expression. Proc. Natl. Acad. Sci. USA *97*, 2075–2080.

Hicks, K.A., and Grossman, A.D. (1996). Altering the level and regulation of the major sigma subunit of RNA polymerase affects gene expression and development in Bacillus subtilus. Mol. Microbiol. *20*, 201–212.

Howard, P., and Maurer, R. (1995). A composite Ets/Pit-1 binding site in the prolactin gene can mediate transcriptional responses to multiple signal transduction pathways. J. Biol. Chem. *270*, 20930–20936.

Huang, C.F., and Ferrell, J.E. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Natl. Acad. Sci. USA *93*, 10078–10083.

Hunter, J., Kassam, A., Winrow, C., Rachubinski, R., and Capone, J. (1996). Crosstalk between the thyroid hormone and peroxisome proliferator-activated receptors in regulating peroxisome proliferator-responsive genes. Mol. Cell. Endocrinol. *116*, 213–221.

Itoh, K., Stevens, B., Schachner, M., and Fields, R.D. (1995). Regulated expression of the neural cell adhesion molecule L1 by specific patterns of neural impulses. Science *270*, 1369–1372.

Karin, M. (1994). Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors. Curr. Opin. Cell Biol. *6*, 415–424.

Keizer, J. (1987). Statistical Thermodynamics of Nonequilibrium Processes (New York: Springer-Verlag).

Keller, A. (1994). Specifying epigenetic states with autoregulatory transcription factors. J. Theor. Biol. *170*, 175–181.

Keller, A. (1995). Model genetic circuits encoding autoregulatory transcription factors. J. Theor. Biol. *172*, 169–185.

Kerszberg, M. (1996). Accurate reading of morphogen concentrations by nuclear receptors: a formal model of complex transduction pathways. J. Theor. Biol. *183*, 95–104.

Kerszberg, M., and Changeux, J. (1994). A model for reading morphogenetic gradients: autocatalysis and competition at the gene level. Proc. Natl. Acad. Sci. USA *91*, 5823–5827.

Kerszberg, M., and Changeux, J. (1998). A simple molecular model of neurulation. Bioessays *20*, 758–770.

Kiebler, M.A., and DesGroseillers, L. (2000). Molecular insights into mRNA transport and local translation in the mammalian nervous system. Neuron *25*, 19–28.

Kientzle, M.J. (1946). Properties of learning curves under varied distributions of practice. J. Exp. Psychol. *36*, 187–211.

King, M.L. (1996). Molecular basis for cytoplasmic localization. Dev. Genet. *19*, 183–189.

Kouzarides, T., and Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. Nature *336*, 646–651.

Lamprecht, R., Hazvi, S., and Dudai, Y. (1997). cAMP response element-binding protein in the amygdala is required for long- but not short-term conditioned taste aversion memory. J. Neurosci. *17*, 8443–8450.

Laurent, M., and Kellershohn, N. (1999). Multistability: a major means of differentiation and evolution in biological systems. Trends Biochem. Sci. *24*, 418–422.

Lee, C., Bae, K., and Edery, I. (1998). The *Drosophila* CLOCK protein undergoes daily rhythms in abundance, phosphorylation, and interactions with the PER-TIM complex. Neuron *21*, 857–867.

Leloup, J.C., and Goldbeter, A. (1998). A model for circadian rhythms in *Drosophila* incorporating the formation of a complex between the PER and TIM proteins. J. Biol. Rhythms *13*, 70–87.

Luo, C., Loros, J.J., and Dunlap, J.C. (1998). Nuclear localization is required for function of the essential clock protein FRQ. EMBO J. *17*, 1228–1235.

Mahaffy, J.M., Jorgensen, D.A., and van der Heyden, R.L. (1992). Oscillations in a model of repression with external control. J. Math. Biol. *30*, 669–691.

Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O., and

Eisenberg, D. (1999). A combined algorithm for genome-wide prediction of protein function. Nature *402*, 83–86.

Martin, K., Michael, D., Rose, J., Barad, M., Casadio, A., Zhu, H., and Kandel, E.R. (1997). MAP kinase translocates into the nucleus of the presynaptic cell and is required for long-term facilitation in *Aplysia*. Neuron *18*, 899–912.

Mayford, M., and Kandel, E.R. (1999). Genetic approaches to memory storage. Trends Genet. *15*, 463–470.

McAdams, H., and Arkin, A. (1997). Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. USA *94*, 814–819.

McAdams, H., and Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. Trends Genet. *15*, 65–69.

McAdams, H., and Shapiro, L. (1995). Circuit simulation of gene networks. Science *269*, 650–656.

Merrow, M.W., Garceau, N., and Dunlap, J.C. (1997). Dissection of a circadian oscillation into discrete domains. Proc. Natl. Acad. Sci. USA *94*, 3877–3882.

Mestl, T., Lemay, C., and Glass, L. (1996). Chaos in high-dimensional neural and gene networks. Physica D *98*, 33–52.

Meyer, T., Waeber, G., Lin, J., Beckmann, W., and Habener, J. (1993). The promoter of the gene encoding cAMP response element binding protein contains cAMP response elements: evidence for positive autoregulation of gene transcription. Endocrinology *132*, 770–780.

Molina, C., Foulkes, N., Lalli, E., and Sassone-Corsi, P. (1993). Inducibility and negative autoregulation of CREM: an alternative promoter directs the expression of ICER, and early response repressor. Cell *75*, 875–886.

O'Leary, F., Byrne, J., and Cleary, L. (1995). Long-term structural remodeling in *Aplysia* sensory neurons requires *de novo* protein synthesis during a critical time period. J. Neurosci. *15*, 3519–3525.

Reppert, S.M. (1998). A clockwork explosion! Neuron *21*, 1–4.

Roenneberg, T., and Merrow, M. (1998). Molecular circadian oscillators: an alternative hypothesis. J. Biol. Rhythms *13*, 167–179.

Ross, I.L., Browne, C.M., and Hume, D. (1994). Transcription of individual genes in eukaryotic cells occurs randomly and infrequently. Immunol. Cell Biol. *72*, 177–185.

Ryu, D.D., and Nam, D.H. (2000). Recent progress in biomolecular engineering. Biotechnol. Prog. *16*, 2–16.

Sabry, J., O'Connor, T., and Kirschner, M.W. (1995). Axonal transport of tubulin in Ti1 pioneer neurons in situ. Neuron *14*, 1247–1256.

Sassone-Corsi, P. (1995). Transcription factors responsive to cAMP. Annu. Rev. Cell Dev. Biol. *11*, 355–377.

Sassone-Corsi, P., Sisson, J.C., and Verma, I.M. (1988). Transcriptional autoregulation of the proto-oncogene *fos*. Nature *334*, 314–319.

Scheper, T., Klinkenberg, D., Pennartz, C., and van Pelt, J. (1999). A mathematical model for the intracellular circadian rhythm generator. J. Neurosci. *19*, 40–47.

Seo, H., Yang, C., Kim, H., and Kim, K. (1996). Multiple protein factors interact with the *cis*-regulatory elements of the proximal promoter in a cell-specific manner and regulate transcription of the dopamine β-hydroxylase gene. J. Neurosci. *16*, 4102–4112.

Sheng, H., Fields, R.D., and Nelson, P. (1993). Specific regulation of immediate-early genes by patterned neuronal activity. J. Neurosci. Res. *35*, 459–467.

Smith, H. (1987a). Oscillations and multiple steady states in a cyclic gene model with repression. J. Math. Biol. *25*, 169–190.

Smith, H. (1987b). Monotone semiflows generated by functional differential equations. J. Diff. Eqs. *66*, 420–442.

Smolen, P., Baxter, D.A., and Byrne, J.H. (1998). Frequency selectivity, multistability, and oscillations emerge from models of gene networks. Am. J. Physiol. *274*, C531–C542.

Smolen, P., Baxter, D.A., and Byrne, J.H. (1999). Effects of macromolecular transport and stochastic fluctuations on the dynamics of gene networks. Am. J. Physiol. *277*, C777–C790.

Smolen, P., Baxter, D.A., and Byrne, J.H. (2000). Modeling transcriptional control in gene networks—methods, recent results, and future directions. Bull. Math. Biol. *62*, 247–292.

Somogyi, R., and Sniegoski, C. (1996). Modeling the complexity of gene networks: understanding multigenic and pleiotropic regulation. Complexity *1*, 45–63.

Somogyi, R., and Sniegoski, C. (1997). The gene expression matrix: towards the extraction of gene network architectures. In Proceedings of the Second World Congress of Nonlinear Analysis (New York: Elsevier Science).

Thayer, M.J., Tapscott, S.J., Davis, R.L., Wright, W.E., Lassar, A.B., and Weintraub, H. (1989). Positive autoregulation of the myogenic determination gene MyoD1. Cell *58*, 241–248.

Thomas, R. (1994). The role of feedback circuits: positive feedback circuits are a necessay condition for positive real eigenvalues of the Jacobian matrix. Ber. Besenges. Phys. Chem. *98*, 1148–1151.

Thomas, R., Thieffry, D., and Kaufman, M. (1995). Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. Bull. Math. Biol. *57*, 247–276.

Tully, T., Preat, T., Boynton, S., and Del Vecchio, M. (1994). Genetic dissection of consolidated memory in *Drosophila melanogaster*. Cell *79*, 35–47.

Walker, W., Fucci, L., and Habener, J. (1995). Expression of the gene encoding transcription factor CREB: regulation by follicle-stimulating hormone-induced cAMP signaling in primary rat sertoli cells. Endocrinology *136*, 3534–3545.

Wen, X.L., Fuhrman, S., Michaels, G., Carr, D., Smith, S., Barker, J., and Somogyi, R. (1998). Large-scale temporal gene expression mapping of central nervous system development. Proc. Natl. Acad. Sci. USA *95*, 334–339.

White, K.P., Rifkin, S.A., Hurban, P., and Hogness, D.S. (1999). Microarray analysis of *Drosophila* development during metamorphosis. Science *286*, 2179–2184.

Winzeler, E.A., Richards, D.R., Conway, A.R., Goldstein, A.L., Kalman, S., McCullough, M.J., McCusker, J.H., Stevens, D.A., Wodicka, L., Lockhart, D.J., and Davis, R.W. (1998). Direct allelic variation scanning of the entire yeast genome. Science *281*, 1194–1197.

Wolf, D.M., and Eeckman, F.H. (1998). On the relationship between genomic regulatory element organization and gene regulatory dynamics. J. Theor. Biol. *195*, 167–186.

Yin, J., and Tully, T. (1996). CREB and the formation of long-term memory. Curr. Opin. Neurobiol. *6*, 264–267.

Yin, J., Del Vecchio, M., Zhou, H., and Tully, T. (1995). CREB as a memory modulator: induced expression of a dCREB2 activator isoform enhances long-term memory in *Drosophila*. Cell *81*, 107–115.

Yisraeli, J.K., Sokol, S., and Melton, D.A. (1990). A two-step model for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of *Vg1* mRNA. Development *108*, 289–298.

Young, M.W. (1998). The molecular control of circadian behavioral rhythms and their entrainment in *Drosophila*. Annu. Rev. Biochem. *67*, 135–152.

Yuh, C.H., Bolouri, H., and Davidson, E.H. (1998). Genomic *cis*-regulatory logic, experimental and computational analysis of a sea urchin gene. Science *279*, 1896–1902.

Zlokarnik, G., Negulescu, P., Knapp, T.E., Mere, L., Burres, N., Feng, L., Whitney, M., Roemer, K., and Tsien, R.Y. (1998). Quantitation of transcription and clonal selection of single living cells with β-lactamase as a reporter. Science *279*, 84–88.