



Analysis of Protein Homeostatic Regulatory Mechanisms in Perturbed Environments at Steady State

CHRISTOPHER SEWELL^{*}, JEFFREY J. MORGAN[†] AND PAUL A. LINDAHL^{*‡§}

^{*}*Department of Chemistry, Texas A&M University, College Station, TX 77843, U.S.A.,*

[†]*Department of Mathematics, Texas A&M University, College Station, TX 77843, U.S.A.*

and ^{‡§}*Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, U.S.A.*

(Received on 20 June 2001, Accepted in revised form on 5 November 2001)

Nine different protein homeostatic regulatory mechanisms were analysed for their ability to maintain a generic protein P within a specified range of a set-point steady-state concentration while perturbed by external processes that altered the rates at which P was produced and/or consumed. Steady state regulatory effectiveness was defined by the area within a rectangular region of “perturbation space”, where axes correspond to rates of positive and negative perturbations. The size of this region differed in accordance with the regulatory elements composing the homeostatic mechanism. Such elements included basic negative feedback control of transcription (in which P, at some high concentration relative to its set-point value, binds to the gene G that encodes it, thereby inhibiting transcription), multiple sequential binding of a feedback effector (two P’s bind sequentially to G), and dimerization of a feedback effector (a P₂ dimer binds to G). Two homeostatic mechanisms included a cascade structure, one with and one without translational feedback control. Another mechanism included feedback control of P degradation. Finally, two mechanisms illustrated the limits of regulatory systems. One lacked all regulatory elements (and included only an invariant rate of P synthesis and degradation) while the other assumed perfect (Boolean) regulation, in which transcription is completely inhibited at $[P] > [P]_{sp}$ and is fully active at $[P] < [P]_{sp}$. All of the systems evaluated are known, but the analytical expressions developed here allow quantitative comparisons between them. These expressions were evaluated at values typical of the average protein in *Escherichia coli*. A method for building regulatory networks by linking semi-independent regulatory modules is discussed.

© 2002 Elsevier Science Ltd. All rights reserved.

Introduction

Regulatory chemical systems consist of interacting components functionally demarcated from other species in surrounding environments. System and environment are not completely isolated, as they share one or more species. Reactions of the

environment perturb the system by altering the concentration of the shared species. The system responds by altering the concentrations of its components in a characteristic manner. The response to such perturbations (or stimuli) defines the system’s behavior. Systems that respond by reestablishing the concentration of a shared component to within a small deviation of a set-point value are called *homeostatic*.

§Author to whom correspondence should be addressed.
E-mail: lindahl@mail.chem.tamu.edu

Living cells use homeostatic regulatory systems to maintain concentrations of cellular components at set-point levels that permit balanced growth and replication despite extracellular and/or intracellular perturbations (Barkai & Leibler, 1997; Becskei & Serrano, 2000). External perturbations include fluctuating nutrient levels, while intracellular perturbations arise from cellular processes other than *core* regulatory processes. For proteins, core processes include biosynthesis (transcription and translation) and degradation or dilution (proteolysis and cell volume changes) (Gottesman & Maurizi, 1992; Axe & Bailey, 1994). Intracellular perturbations include those arising from the physiological function of the particular protein. Proteins involved in cell-cycle processes tend to change concentration when they are activated (e.g. the concentration of their unbound state would tend to decline upon assembly of a structure for which the protein is a component) and change again in the opposite direction when deactivated (e.g. when the structure disassembles). Proteins involved in metabolism or gene expression tend to change concentration as they bind or unbind substrates, products, effector molecules or other proteins, or as they are covalently modified or unmodified.

Cells contain large numbers of different types of proteins, each presumably regulated by homeostatic mechanisms that control gene expression. Since proteins interact with each other, homeostatic systems perturb, and are perturbed by, each other. Numerous strategies have been developed to understand and model such regulatory networks, and tremendous progress has been made over the past 30 years (Rosen, 1968; Glass & Kauffman, 1973; Glass, 1975; Thomas, 1973, 1991; Tyson & Othmer, 1978; Meiske & Reich, 1987; Hlavacek & Savageau, 1996; Mestl *et al.*, 1995; Omholt *et al.*, 1998; McAdams & Arkin, 1998; McAdams & Shapiro, 1995; Ni & Savageau, 1996; Hofmeyr & Cornish-Bowdon, 1996; Savageau, 1998; Wolf & Eeckman, 1998; Chen *et al.*, 1999; Edwards & Glass, 2000; Hofmeyr & Westerhoff, 2001). Excellent reviews and books have recently appeared (Smolen *et al.*, 2000a, b; Bower & Bolouri, 2001). One approach involves constructing networks of genes that are either “on” or “off” according to defined

logical/Boolean relationships, while another involves constructing chemical mechanisms and numerically analysing the corresponding differential equations. The Boolean approach is less realistic but more computationally tractable and thus applicable for modeling highly complex networks. The mechanistic/differential equations approach is more realistic and better for simpler systems where reliable kinetic data (rate constants and concentrations) are available. Unfortunately, such data are generally *not* available.

For proteins with low cellular copy numbers, dynamics are simulated better using stochastic rather than deterministic assumptions (McAdams & Arkin, 1997, 1999). However, stochastic systems are more difficult to solve, and systems solved deterministically can be viewed loosely (and with caution) as reflecting the average of an ensemble of cells (Axe & Bailey, 1994).

Despite the progress made in understanding regulatory networks, further advances are required before global networks operating in even the simplest living systems can be constructed from genomic, proteomic, transcriptomic and metabolomic data (Laub *et al.*, 2000). We have constructed numerous simple homeostatic mechanisms for regulating a generic protein P. In this study, we analyse the magnitude of perturbations that these core systems can tolerate before steady-state concentrations of P shift beyond a specified range centered about an unperturbed set-point concentration $[P]_{sp}$. The regulatory effectiveness of each system, thus defined, was quantified and correlated to the regulatory *elements* of which they were

|| Abbreviations. RM1–RM9, regulatory mechanisms 1–9 as illustrated in Fig. 1 and defined in the text; P, a generic protein with properties of the average protein in *Escherichia coli*; PP, dimer of P; G, the gene encoding P, presumed to catalyse the synthesis of P directly in RM1–RM6 and RM9; GP, GPP, GPPP, GPPPP, complexes with G and 1–4 P's; M, the mRNA transcript encoding P in RM7 and RM8; ϵ , a unitless factor indicating regulatory “tightness”, with values ranging from 0 to 1; $[X]$, concentration of the designated species ($X=P, G, \text{etc.}$) under unspecified conditions; $[P]_{sp}$, the set-point concentration of P; $[P]_{max}$, the maximum $[P]$ in which a system is regulated; $[P]_{min}$, the minimum $[P]$ in which a system is regulated; $[G]_{sp}$, $[G]$ when $[P] = [P]_{sp}$; $[G]_{min}$, $[G]$ when $[P] = [P]_{min}$; $[G]_{max}$, $[G]$ when $[P] = [P]_{max}$; $[G]_{tot}$, the total concentration of G, presumed to be 1 nM.

composed. Although the systems analysed are well known, the analytical expressions presented allow the effectiveness of these systems to be quantitatively compared. This approach suggests a strategy for building regulatory networks.

Analysis and Results

METHODOLOGY

Regulatory effectiveness in homeostatic mechanisms can be divided into steady-state and dynamic contributions. Steady state effectiveness refers to the ability to maintain steady-state $[P]$ within a small deviation of $[P]_{sp}$ while being perturbed by synthesis and/or degradation processes external to core processes. Dynamic effectiveness refers to the *rate* of recovery from a perturbation in $[P]$. Along these lines, Axe & Bailey (1994) have discussed a “step/response” approach to analyse steady state effectiveness and an “impulse/response” approach to analyse dynamic effectiveness.

We constructed a series of nine simple homeostatic mechanisms (RM1–RM9) and examined their steady state regulatory effectiveness using Maple 6 (Waterloo Maple). Each consisted of a set of core reactions (Fig. 1) augmented by two opposing perturbation reactions. Since the mechanisms were examined at steady state, the ordinary differential equations describing the dynamics of these systems were equated to zero and the resulting algebraic expressions, as well as mass conservation relationships and chemical equilibrium expressions (involving thermodynamic equilibrium constants K_{eq}) were analysed. In all but two mechanisms (RM7 and RM8), the gene encoding P (called G) catalyses the synthesis of P in a first-order process (with rate constant k_1). Numerous cellular components affect the rate of protein biosynthesis; however, in these analyses the concentrations of all such components are presumed to be constant and are incorporated into k_1 . In all cases except RM9, the degradation of P was assumed to be a first-order process dependent on P (with rate constant k_2). The perturbation reactions represent a compilation of external processes which increase and decrease $[P]$, with overall rates k_3 and $k_4[P]$, respectively. For RM1–RM6, at steady state and

in the presence of perturbation reactions,

$$[P] = \frac{k_1[G] + k_3}{k_2 + k_4}. \quad (1)$$

In the absence of perturbations,

$$[P]_{sp} = \frac{k_1[G]_{sp}}{k_2}. \quad (2)$$

The fraction of $[G]_{tot}$ in the catalytically active form $[G]$ varies for each mechanism, typically as a function of $[P]$ and of one or more K_{eq} 's. Plots of these functions ($[G]/[G]_{tot}$ vs. $\log[P]$) for each mechanism are shown in Fig. 2. Maximum sensitivity (percent change in $[G]/[G]_{tot}$ caused by a percent change in $[P]$) occurs at $[G]/[G]_{tot} = 1/2$ (Goldbeter & Koshland, 1982). Thus, K_{eq} 's were chosen such that $[P] = [P]_{sp}$ when $[G]/[G]_{tot} = 1/2$, as this should allow the maximum regulatory ability in the presence of perturbations. Comparable analyses for those mechanisms (RM7 and RM8) in which M catalysed the synthesis of P and the one mechanism in which GP catalysed the degradation of P (RM9) differed slightly, as described below. For $[G]_{tot} = 1$ nM, eqn (2) becomes

$$[P]_{sp} = \frac{k_1}{2k_2}. \quad (3)$$

LIMITS OF REGULATORY EFFECTIVENESS

Our approach evaluates the relative perturbation magnitude that a system can tolerate, rather than its *sensitivity* to an incremental change in the stimulus (Koshland *et al.*, 1982; Goldbeter & Koshland, 1982; Kholodenko *et al.*, 1997). We define steady state regulatory effectiveness by (4)

$$R_{ss} = \frac{[P] - [P]_{sp}}{[P]_{sp}}, \quad (4)$$

where R_{ss} indicates the relative difference between steady state perturbed $[P]$ and unperturbed set-point $[P]_{sp}$ concentrations. We introduce a “regulatory tightness” parameter ε ranging from 0 to 1. Mechanisms operating so that $-\varepsilon \leq R_{ss} \leq +\varepsilon$ were considered to be regulated. Expression (5) gives the maximum and minimum $[P]$ at the limits of effective regulation,

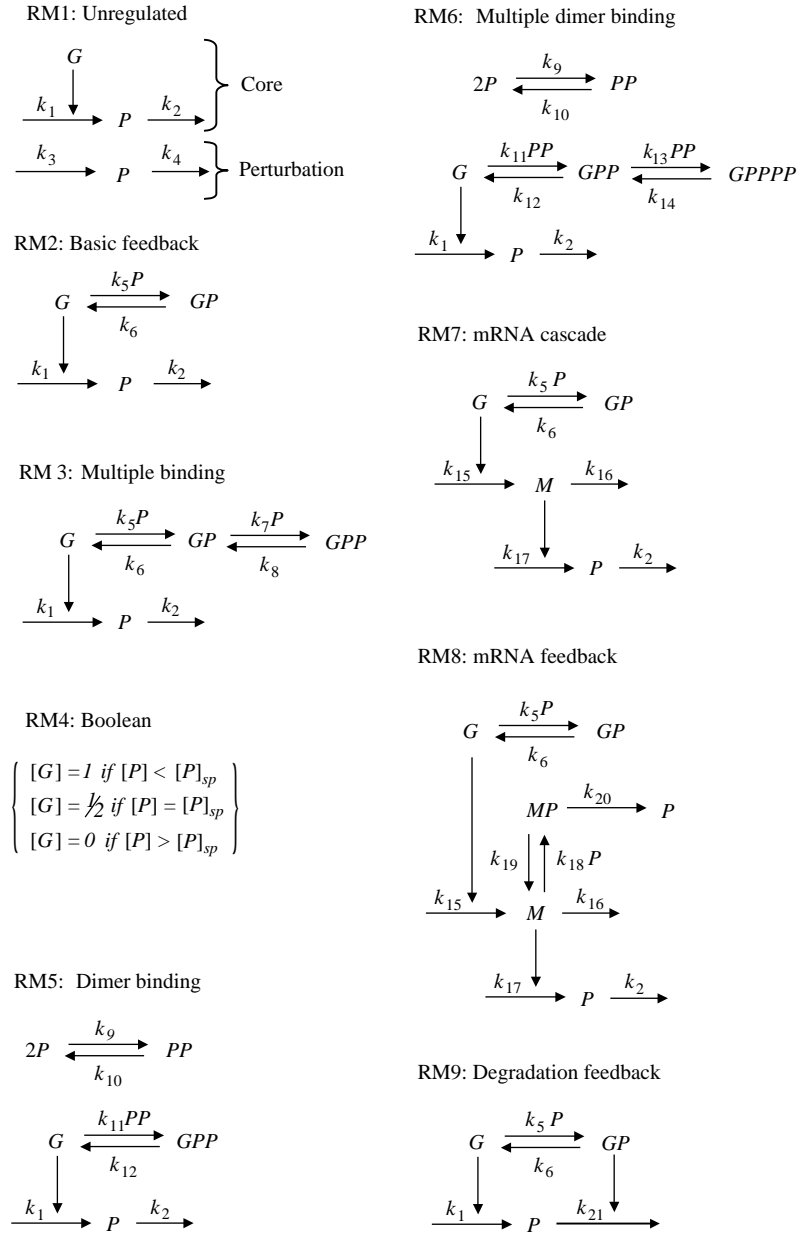


FIG. 1. Schematic representations of the nine regulatory mechanisms examined. Horizontal arrows indicate reactions. Absence of a substrate or product in a reaction indicates an assumed concentration-invariance for the species. Vertical arrows indicate catalytic steps, with the catalyst located at the origin of the arrow and the reaction catalysed at the terminus. The perturbation reactions shown for RM1 also apply to RM2–RM9.

respectively.

$$\begin{aligned} [P]_{max} &= (1 + \varepsilon)[P]_{sp}, \\ [P]_{min} &= (1 - \varepsilon)[P]_{sp}. \end{aligned} \quad (5)$$

As tighter regulation is required, ε is set to values approaching zero, and the difference between $[P]_{max}$ and $[P]_{min}$ vanishes.

The concentration of G at $[P]_{min}$ (called $[G]_{max}$) was found by substituting $[P]_{min}$ for $[P]$ into $[G]/[G]_{tot}$ functions that had K_{eq} 's optimized for maximum sensitivity at $[P]_{sp}$. Likewise, $[G]_{min}$ was found by substituting $[P]_{max}$ for $[P]$. The resulting $[P]_{max} : G_{min}$ and $[P]_{min} : G_{max}$ pairs were substituted into eqn (1), and resulting expressions were solved for k_3 , yielding eqns (6) and (7). These equations describe regulatory

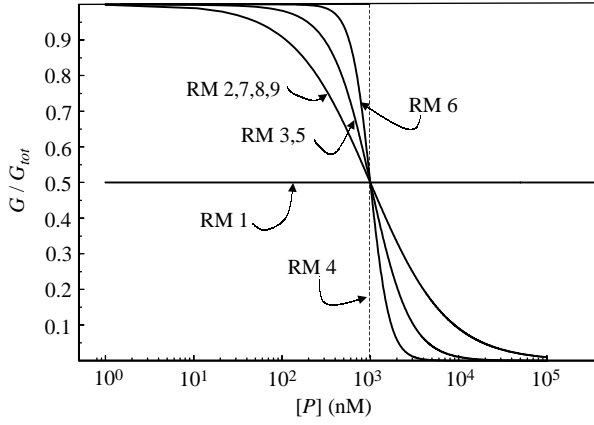


FIG. 2. The proportion of gene G that is unbound and catalytically active as a function of the concentration of generic protein P for each mechanism. In all cases, $[P]_{sp} = 1000$ nM.

boundary lines in a $k_3 : k_4$ perturbation space.

$$k_3 = \frac{1}{2}k_1(1 + \varepsilon - 2G_{min}) + (1 + \varepsilon)[P]_{sp}k_4, \quad (6)$$

$$k_3 = \frac{1}{2}k_1(1 - \varepsilon - 2G_{max}) + (1 - \varepsilon)[P]_{sp}k_4. \quad (7)$$

By our definition, systems operating within these lines are effectively regulated (Fig. 3). Since k_3 and k_4 are opposing processes, the presence of both diminishes the perturbing effect of either acting alone, which renders systems regulated at extremely high k_3 and k_4 combinations. Of greater interest are situations in which either k_3 or k_4 may be present in the absence of the other. The region of perturbation space in which systems are regulated under these “unsynchronized” conditions is indicated by the shaded rectangle of Fig. 3. The magnitude of this area (A_{unsync}) will be used to quantify the effectiveness of regulatory mechanisms (Fig. 4). $A_{unsync} = k_{3max} \times k_{4max}$, where k_{3max} is the maximum k_3 for which systems are regulated when $k_4 = 0$. Likewise, k_{4max} is the maximum k_4 for which systems are regulated when $k_3 = 0$. The analytical expressions for these parameters, assuming $[P]_{sp}$ as given in eqn (3), are

$$k_{3max} = \frac{1}{2}k_1(1 + \varepsilon - 2G_{min}), \quad (8)$$

$$k_{4max} = \frac{k_2(2G_{max} + \varepsilon - 1)}{(1 - \varepsilon)}. \quad (9)$$

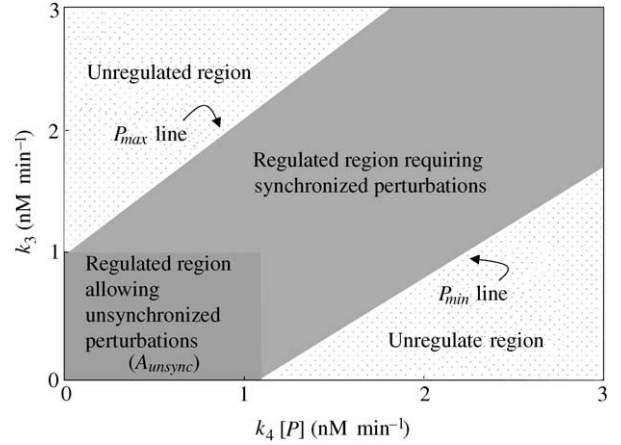


FIG. 3. Plot of two-dimensional “perturbation space” (for RM1). P_{max} and P_{min} lines represent $k_3 : k_4$ combinations at which the perturbed steady-state concentration of P deviates from the unperturbed set-point concentration ($[P]_{sp}$) by a factor of $(1 \pm \varepsilon)$, respectively. Lines are drawn for RM1, with $k_1 = 20$ min^{-1} , $k_2 = 0.01$ min^{-1} and $\varepsilon = 0.1$. Systems with $k_3 : k_4$ combinations located between these lines are regulated while those with combinations located outside of these lines are not. The regulated region is subdivided into “synchronized” and “unsynchronized” areas. Since k_3 and k_4 processes oppose each other, the simultaneous presence of both (indicated by the synchronized region) attenuates their perturbing effects. The unsynchronized rectangular area (A_{unsync}) best reflects regulatory effectiveness, as systems within it are regulated in the presence of either k_3 or k_4 perturbations.

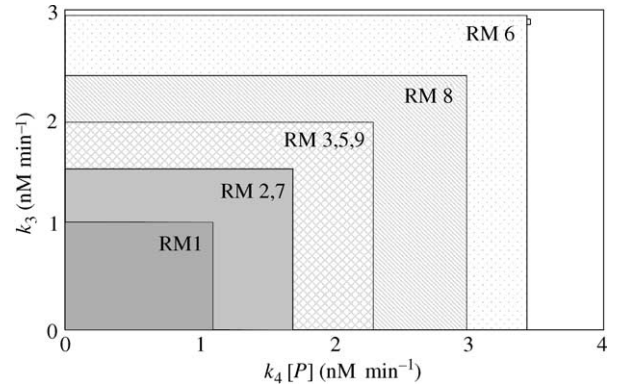


FIG. 4. Plots of the unsynchronized rectangular areas for various regulatory mechanisms. Intercepts and areas (A_{unsync}) of these rectangles are given in Table 1. The corresponding rectangle for RM4 (not shown) would intercept the ordinate and abscissa at 11.0 and 12.2, respectively.

NUMERICAL CONSIDERATIONS

Evaluating the analytical expressions of Table 1 at specific core rate constants and ε values allowed the regulatory mechanisms to be quantitatively compared. In our examples, ε was

TABLE 1

	k_{3max} (nM min ⁻¹)	k_{4max} (min ⁻¹)	$A_{unsync} \times [P]_{sp}$ (nM ² min ⁻²)
RM1	$\frac{1}{2}k_1\varepsilon$	$\frac{k_2\varepsilon}{(1-\varepsilon)}$	1.11
RM2	$\frac{1}{2}\frac{k_1(3+\varepsilon)\varepsilon}{(2+\varepsilon)}$	$\frac{k_2(3-\varepsilon)\varepsilon}{(1-\varepsilon)(2-\varepsilon)}$	2.50
RM3	$\frac{1}{2}\frac{k_1(\varepsilon^2+3\varepsilon+4)\varepsilon}{(\varepsilon^2+2\varepsilon+2)}$	$\frac{k_2(\varepsilon^2-3\varepsilon+4)\varepsilon}{(1-\varepsilon)(\varepsilon^2-2\varepsilon+2)}$	4.44
RM4	$\frac{1}{2}k_1(1+\varepsilon)$	$\frac{k_2(1+\varepsilon)}{(1-\varepsilon)}$	134.00
RM5	$\frac{1}{2}\frac{k_1(\varepsilon^2+3\varepsilon+4)\varepsilon}{(\varepsilon^2+2\varepsilon+2)}$	$\frac{k_2(\varepsilon^2-3\varepsilon+4)\varepsilon}{(1-\varepsilon)(\varepsilon^2-2\varepsilon+2)}$	4.44
RM6	$\frac{1}{2}\frac{k_1(6+10\varepsilon+10\varepsilon^2+5\varepsilon^3+\varepsilon^4)\varepsilon}{(2+4\varepsilon+6\varepsilon^2+4\varepsilon^3+\varepsilon^4)}$	$\frac{k_2(6-10\varepsilon+10\varepsilon^2-5\varepsilon^3+\varepsilon^4)\varepsilon}{(1-\varepsilon)(2-4\varepsilon+6\varepsilon^2-4\varepsilon^3+\varepsilon^4)}$	9.86
RM7	$\frac{1}{2}\frac{k_a(3+\varepsilon)\varepsilon}{(2+\varepsilon)}$	$\frac{k_2(3-\varepsilon)\varepsilon}{(1-\varepsilon)(2-\varepsilon)}$	2.50
RM8	$\frac{1}{2}\frac{k_b(\varepsilon^2+4\varepsilon+5)\varepsilon}{(2+3\varepsilon+\varepsilon^2)}$	$\frac{k_2(\varepsilon^2-4\varepsilon+5)\varepsilon}{(1-\varepsilon)(\varepsilon^2-3\varepsilon+2)}$	7.02
RM9	$k_1\varepsilon$	$\frac{k_{21}\varepsilon}{(1-\varepsilon)}$	4.44

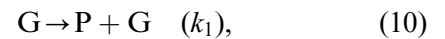
Note: A_{unsync} evaluated with $\varepsilon = 0.1$, $k_1 = 20 \text{ min}^{-1}$, $k_2 = 0.01 \text{ min}^{-1}$, $k_{15} = 4 \text{ min}^{-1}$, $k_{16} = 1 \text{ min}^{-1}$, $k_{17} = 5 \text{ min}^{-1}$, $k_{18} = 0.002 \text{ nM}^{-1} \text{ min}^{-1}$, $k_{19} = 1 \text{ min}^{-1}$, $k_{20} = 1 \text{ min}^{-1}$, and $k_{21} = 0.02 \text{ min}^{-1}$. In RM7, $k_a = k_{15}k_{17}/k_{16} = 20 \text{ min}^{-1}$. In RM8, $k_b = \sqrt{2k_2k_{15}k_{17}(k_{19} + k_{20})}/k_{18}k_{20} = 20 \text{ min}^{-1}$.

assigned the value 0.1 while other parameters were based on those relevant to the average protein in *Escherichia coli* (Neidhardt & Umbarger, 1996; Hargrove, 1994; Voet & Voet, 1995; Lewis, 1998). Consider a hypothetical newborn prokaryotic cell with a cytoplasmic volume of $1.66 \times 10^{-15} \text{ l}$ containing 1000 copies (= 1000 nM) of a generic 40 kDa protein P and 1 copy of its gene G (in an active form half of the time). Using this volume, the copy number of any component equals its concentration in nM. In the absence of perturbations, P is synthesized at a rate of 10 nM min^{-1} . This process is catalysed by G with an apparent first-order rate constant $k_1 = 20 \text{ min}^{-1}$ (i.e. $10 \text{ nM min}^{-1}/0.5 \text{ nM}$). P is stable to protease degradation. However, during each 50 min cell-cycle, cell volume doubles and P is diluted at an average rate of 10 nM min^{-1} . The rate of cell growth (and thus the ‘‘degradation’’ rate) depends on P, yielding an apparent first-order degradation rate-constant of $k_2 = 0.01 \text{ min}^{-1}$ (i.e. $10 \text{ nM min}^{-1}/1000 \text{ nM}$). At steady state,

eqn (3) indicates that $[P]_{sp} = 1000 \text{ nM}$ ($20 \text{ min}^{-1} \times 0.5 \text{ nM}/0.01 \text{ min}^{-1}$).

REGULATORY MECHANISM ONE (RM1, ‘‘UNREGULATED’’)

RM1 lacks regulatory elements and is the control mechanism against which all others are compared. It is composed only of a G-dependent process that synthesizes P and a P-dependent one that degrades P (Fig. 1). The specific reactions and rate-constants defining RM1 (core as well as perturbation reactions) include the following:



The corresponding differential equation that describes the time-dependent change in $[P]$ is

$$d[P]/dt = k_1[G] - k_2[P] + k_3 - k_4[P]. \quad (14)$$

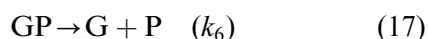
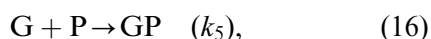
At steady state,

$$k_1[G] + k_3 = k_2[P] + k_4[P], \quad (15)$$

which rearranges to eqn (1). In RM1, $[G]/[G]_{tot}$ is a horizontal line (slope=0) at $[G]_{sp} = [G]_{min} = [G]_{max} = 1/2$. (The value 1/2 is chosen for easy comparison to other mechanisms.) Substituting these values into eqns (8) and (9) yields the k_{3max} , k_{4max} , and A_{unsync} values for RM1 in Table 1. A_{unsync} is directly proportional to the core synthesis and degradation rates k_1 and k_2 , indicating that regulation improves when P is synthesized and degraded rapidly.

REGULATORY MECHANISM TWO (RM2,
"BASIC NEGATIVE FEEDBACK")

RM2 is similar to RM1, except that it includes negative feedback control of transcription (P binds G, affording catalytically inactive GP) (Fig. 1). This well-known regulatory element has been extensively studied (e.g. Thomas *et al.*, 1995). Defining reactions and rate constants for RM2 include eqns (10)–(13) and



as well as conservation relationship:

$$[G]_{tot} = 1 \text{ nM} = [G] + [GP]. \quad (18)$$

The corresponding differential equations include

$$\begin{aligned} d[P]/dt = k_1[G] - k_2[P] + k_3 - k_4[P] \\ - k_5[G][P] + k_6[GP], \end{aligned} \quad (19)$$

$$d[GP]/dt = -d[G]/dt = k_5[G][P] - k_6[GP]. \quad (20)$$

At steady state, eqn (19) rearranges into eqn (15) [and alternatively eqn (1)] since eqn (20) becomes

$$k_5[G][P] = k_6[GP] \quad (21)$$

which rearranges into equilibrium expression

$$K_{GP} = \frac{k_5}{k_6} = \frac{[GP]}{[G][P]}, \quad (22)$$

$[GP]$ can be eliminated using eqn (18), yielding the $[G]/[G]_{tot}$ function

$$[G]/[G]_{tot} = \frac{1}{K_{GP}[P] + 1}. \quad (23)$$

The K_{GP} relationship yielding maximum sensitivity occurs when $[G]/[G]_{tot} = 1/2$. Solving eqn (23) for K_{GP} under this condition yields

$$K_{GP} = \frac{1}{[P]_{sp}} = \frac{2k_2}{k_1}. \quad (24)$$

$[P]_{min}$ and $[P]_{max}$ as defined in eqn (5) were substituted with eqn (24) into eqn (23), yielding

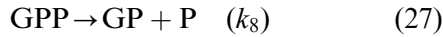
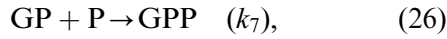
$$\begin{aligned} G_{min} &= \frac{1}{2 + \varepsilon}, \\ G_{max} &= \frac{1}{2 - \varepsilon}. \end{aligned} \quad (25)$$

k_{3max} , k_{4max} and A_{unsync} values (Table 1) were obtained by substituting eqn (25) into eqns (8) and (9). A_{unsync} for RM2 is larger than that for RM1 by a factor of 2.3. Thus, including basic negative feedback control of transcription *more than doubles* steady-state regulatory effectiveness. This improvement arises because the binding of P to G prevents G from catalysing the synthesis of more P. Synthesis of P is inhibited when $[P] > [P]_{sp}$ and stimulated when $[P] < [P]_{sp}$, as the proportion of $[G]/[G]_{tot}$ changes. In contrast, the proportion of $[G]_{tot}$ available for P catalysis in RM1 is invariant to changes in $[P]$. Although the autoregulated and unregulated homeostatic systems genetically engineered by Becskei & Serrano (2000) are only qualitatively similar to RM2 and RM1, respectively, their autoregulated system was also approximately twice as stable as their unregulated one.

REGULATORY MECHANISM THREE (RM3, "MULTIPLE
BINDING FEEDBACK")

Multiple copies of autoregulated proteins and transcription factors often bind to promoter sites on the DNA that control gene expression. These

binding events are typically cooperative, with the equilibrium constant for the first event (K_{GP}) being smaller (i.e. weaker) than those of subsequent events. Ptashne (1992) has shown that binding of multiple λ Repressor transcriptional factors increases the sensitivity of the transcriptional “switch” to changes in Repressor concentration. RM3 (Fig. 1) was analysed to determine the effect of multiple sequential binding on homeostatic regulatory effectiveness. RM3 is the same as RM2 except that two P’s bind sequentially to G (affording states GP and GPP, respectively). The defining reactions and rate constants for RM3 include eqns (10)–(13), (16), (17) and



as well as conservation relationship:

$$[G]_{tot} = 1 \text{ nM} = [G] + [GP] + [GPP]. \quad (28)$$

Corresponding differential equations include

$$d[P]/dt = k_1[G] - k_2[P] + k_3 - k_4[P] - k_5[G][P] + k_6[GP] - k_7[P][GP] + k_8[GPP], \quad (29)$$

$$d[G]/dt = -k_5[G][P] + k_6[GP], \quad (30)$$

$$d[GP]/dt = k_5[G][P] - k_6[GP] - k_7[GP][P] + k_8[GPP], \quad (31)$$

$$d[GPP]/dt = k_7[GP][P] - k_8[GPP]. \quad (32)$$

At steady state, expressions (15), (21) and (33) are obtained.

$$k_7[GP][P] = k_8[GPP]. \quad (33)$$

Rearrangement yields equilibrium expressions (22) and

$$K_{GPP} = \frac{k_7}{k_8} = \frac{[GPP]}{[GP][P]}. \quad (34)$$

Substituting eqns (22) and (34) into eqn (28) yields

$$[G]/[G]_{tot} = \frac{1}{1 + K_{GP}[P] + K_{GPP}K_{GP}[P]^2}. \quad (35)$$

For maximum sensitivity at $[P]_{sp}$,

$$K_{GP} = \frac{1}{[P]_{sp}(K_{GPP}[P]_{sp} + 1)}. \quad (36)$$

Substituting relationship (36) and (5) into eqn (35) yields

$$G_{min} = \frac{K_{GPP}k_1 + 2k_2}{2K_{GPP}k_1 + 4k_2 + 2k_2\varepsilon + 2K_{GPP}k_1\varepsilon + K_{GPP}k_1\varepsilon^2},$$

$$G_{max} = \frac{K_{GPP}k_1 + 2k_2}{2K_{GPP}k_1 + 4k_2 - 2k_2\varepsilon - 2K_{GPP}k_1\varepsilon + K_{GPP}k_1\varepsilon^2}. \quad (37)$$

When K_{GPP} becomes exceedingly large,

$$G_{min} = \frac{1}{2 + 2\varepsilon + \varepsilon^2},$$

$$G_{max} = \frac{1}{2 - 2\varepsilon + \varepsilon^2}. \quad (38)$$

These limiting relationships for G_{min} and G_{max} were substituted into eqns (8) and (9) to yield the k_{3max} , k_{4max} , and A_{unsync} values given in Table 1. Relative to RM2, A_{unsync} is larger by a factor of 1.8. This improvement reflects an idealized case since the analysis assumes an exceedingly large value for K_{GPP} .

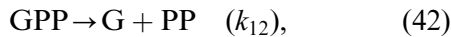
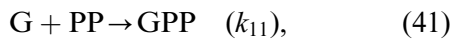
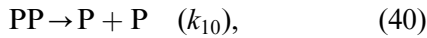
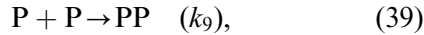
REGULATORY MECHANISM FOUR (RM4, “BOOLEAN NEGATIVE FEEDBACK”)

RM3 regulates better than RM2 because the binding of 2 P’s to G is cooperative and this increases the sensitivity of negative feedback (making it ultrasensitive). If even more P’s were allowed to bind, sensitivity would increase further, yielding in the limit a discontinuous function $[G]/[G]_{tot} = 1$ at $[P] < [P]_{sp}$ and $[G]/[G]_{tot} = 0$ at $[P] > [P]_{sp}$. We may also specify $[G]/[G]_{tot} = 1/2$ at $[P] = [P]_{sp}$. This idealized, best-case sensitivity is illustrated in Fig. 2 by Boolean mechanism RM4. The reactions defining RM4 are eqns (10)–(13), yielding differential equation (14), steady-state relationship (15) and conservation relationship (18). The only significant difference between RM4 and RM2 is the function $[G]/[G]_{tot}$. In RM4, $[G]_{min} = 0$, $[G]_{max} = 1$, and $[G]_{sp} = 1/2$. Substituting these values into

eqns (8) and (9) yields the expressions for k_{3max} , k_{4max} , and A_{unsync} in Table 1. A_{unsync} is more than 50 times larger than that for RM2 and over 100 times better than that for RM1! This reveals the advantage of “switch-like” genetic responses (Ferrell, 1996–1998) in homeostatic regulation. However, it also highlights the unrealistic improvement in regulation achieved when Boolean systems are substituted for simple chemical mechanistic feedback systems in hypothetical genetic networks.

REGULATORY MECHANISM FIVE
(RM5, “DIMERIZATION”)

Most if not all transcription factors exist in monomer \rightleftharpoons oligomer equilibria, with dimers and tetramers being the most common oligomeric forms. These DNA-binding proteins typically bind G’s as oligomers, increasing the sensitivity of gene expression to changes in the concentration of the regulated protein (Ptashne, 1992). Regulatory mechanism RM5 (Fig. 1) was constructed and analysed to quantify the degree to which this arrangement improves homeostatic regulation. Defining reactions, rate constants, and conservation relationship for RM5 include eqns (10)–(13) and



$$[G]_{tot} = 1 \text{ nM} = [G] + [GPP]. \quad (43)$$

Corresponding differential equations include

$$\begin{aligned} d[P]/dt = & k_1[G] - k_2[P] + k_3 - k_4[P] \\ & - 2k_9[P]^2 + 2k_{10}[PP], \end{aligned} \quad (44)$$

$$d[G]/dt = -k_{11}[G][PP] + k_{12}[GPP], \quad (45)$$

$$\begin{aligned} d[PP]/dt = & k_9[P]^2 - k_{10}[PP] \\ & - k_{11}[G][PP] + k_{12}[GPP], \end{aligned} \quad (46)$$

$$d[GPP]/dt = k_{11}[G][PP] - k_{12}[GPP]. \quad (47)$$

Steady-state equations include eqn (15) and

$$k_9[P]^2 = k_{10}[PP], \quad (48)$$

$$k_{11}[G][PP] = k_{12}[GPP]. \quad (49)$$

Equilibrium relationships include

$$K_{PP} = \frac{k_9}{k_{10}} = \frac{[PP]}{[P]^2}, \quad (50)$$

$$K_{GP2} = \frac{k_{11}}{k_{12}} = \frac{[GPP]}{[G][PP]}. \quad (51)$$

Equations (50) and (51) were substituted into eqn (43) yielding

$$[G]/[G]_{tot} = \frac{1}{1 + K_{PP}K_{GP2}[P]^2}. \quad (52)$$

For maximum sensitivity at $[P]_{sp}$,

$$K_{PP} = \frac{1}{K_{GP2}[P]_{sp}^2} = \frac{4k_2^2}{K_{GP2}k_1^2}. \quad (53)$$

This relationship and eqn (5) were substituted back into eqn (52) to yield $[G]_{max}$ and $[G]_{min}$ values [eqn (38)] that were substituted into eqns (8) and (9) to yield the k_{3max} , k_{4max} , and A_{unsync} for this mechanism. These values (Table 1) are identical to those obtained for RM3 under optimum circumstances. Thus, regulatory systems in which monomeric transcription factors dimerize and then bind genetic loci are equally effective in homeostatic regulation as those in which two copies of a monomeric transcription factor bind sequentially (assuming an exceedingly large second binding constant). Given finite binding constants, dimerization mechanisms will be more effective than otherwise equivalent multiple sequential binding mechanisms. These relationships arise because the presence of the intermediate GP form in RM3 (absent in RM5) detracts from the regulatory ability of the mechanism. Optimal behavior is achieved when K_{GP} and K_{GPP} are adjusted to eliminate $[GP]$. In the limiting case when $[GP] = 0$, the two mechanisms become identical.

REGULATORY MECHANISM SIX (RM6,
"TWO-DIMER BINDING")

RM6 includes the dimerization of P and the sequential binding of two dimers to G (Fig. 1). It was analysed to assess whether combining these two regulatory elements would improve regulatory effectiveness. Defining reactions, rate constants, and conservation relationship for RM6 include eqns (10)–(13), (39)–(42) and



$$[G]_{\text{tot}} = 1 \text{ nM} = [G] + [GPP] + [GPPPP]. \quad (56)$$

Corresponding differential equations include eqns (44), (45) and

$$\begin{aligned} d[PP]/dt = & k_9[P]^2 - k_{10}[PP] - k_{11}[G][PP] \\ & + k_{12}[GPP] - k_{13}[GPP][PP] + k_{14}[GPPPP], \end{aligned} \quad (57)$$

$$\begin{aligned} d[GPP]/dt = & k_{11}[G][PP] - k_{12}[GPP] \\ & - k_{13}[PP][GPP] + k_{14}[GPPPP], \end{aligned} \quad (58)$$

$$d[GPPPP]/dt = k_{13}[PP][GPP] - k_{14}[GPPPP]. \quad (59)$$

At steady state, eqns (15), (48), (49) and (60) are obtained.

$$k_{13}[PP][GPP] = k_{14}[GPPPP]. \quad (60)$$

Rearrangement yields equilibrium expressions (50), (51) and

$$K_{GP4} = \frac{k_{13}}{k_{14}} = \frac{[GPPPP]}{[GPP][PP]}. \quad (61)$$

Substituting eqns (50), (51) and (61) into eqn (56) yields

$$[G]/[G]_{\text{tot}} = \frac{1}{1 + K_{P2}K_{GP2}[P]^2 + K_{P2}^2K_{GP2}K_{GP4}[P]^4}. \quad (62)$$

For maximum sensitivity at $[P]_{sp}$,

$$K_{GP2} = \frac{1}{K_{P2}[P]_{sp}^2(1 + K_{GP4}K_{P2}[P]_{sp}^2)}. \quad (63)$$

This equation and eqn (5) were substituted back into eqn (62) to obtain $[G]_{\text{min}}$ and $[G]_{\text{max}}$. At exceedingly large values of K_{GP4} ,

$$\begin{aligned} G_{\text{min}} &= \frac{1}{2 + 4\varepsilon + 6\varepsilon^2 + 4\varepsilon^3 + \varepsilon^4}, \\ G_{\text{max}} &= \frac{1}{2 - 4\varepsilon + 6\varepsilon^2 - 4\varepsilon^3 + \varepsilon^4}. \end{aligned} \quad (64)$$

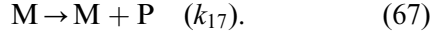
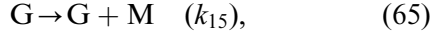
These values were substituted into eqns (8) and (9) to yield $k_{3\text{max}}$, $k_{4\text{max}}$, and A_{unsync} values for this mechanism (Table 1). A_{unsync} was 2.2 times greater than that for either RM3 or RM5. Thus, regulatory effectiveness improves significantly by combining transcription-factor dimerization and multiple sequential binding. This situation is found with the λ Repressor, which dimerizes and then binds in multiple copies to its operator (Ptashne, 1992).

REGULATORY MECHANISM SEVEN (RM7,
"TRANSCRIPTION/TRANSLATION CASCADE")

Transcription and translation may be viewed as a two-enzyme (E_1 , E_2) cascade. In such arrangements, a signal binds E_1 , converting it to a form that activates E_2 to catalyse a reaction $S \rightarrow P$. In cascades, the rate of the ultimate reaction is amplified relative to the situation in which the signal binds E_2 directly (Chock *et al.*, 1980; Shacter-Noiman *et al.*, 1983; Shacter *et al.*, 1984). The amplification factor depends on the number of levels in the cascade and the rates of catalysis at each level.

In transcription/translation regulatory cascades, all of the components required for transcription, including G, together constitute E_1 and catalyse the synthesis of mRNA (M) from nucleotides. All components required for translation, including M, together constitute E_2 , which catalyses the synthesis of P from amino acids. The only major difference between transcription/translation cascades and "standard" cascades is that the product of the E_2 reaction (P) is also the signal, affording feedback behavior. In RM7 (Fig. 1), we assess the degree

to which this arrangement affects homeostatic regulatory ability. The defining reactions and conservation relationship for RM7 are eqns (11)–(13), (16)–(18) and (65)–(67).



Corresponding differential equations are eqn (20) and

$$\begin{aligned} d[P]/dt = & k_{17}[M] - k_2[P] + k_3 - k_4[P] \\ & - k_5[G][P] + k_6[GP], \end{aligned} \quad (68)$$

$$d[M]/dt = k_{15}[G] - k_{16}[M]. \quad (69)$$

The corresponding steady-state equations are eqn (21) and

$$k_{17}[M] + k_3 = k_2[P] + k_4[P], \quad (70)$$

$$k_{15}[G] = k_{16}[M]. \quad (71)$$

Equilibrium expression (22) applies as does the $[G]/[G]_{tot}$ function (23). Equations (70) and (71) can be rearranged to give

$$[P] = \frac{k_{17}[M] + k_3}{k_2 + k_4} = \frac{k_{17}k_{15}[G] + k_3k_{16}}{k_2k_{16} + k_4k_{16}}. \quad (72)$$

This equation replaces eqn (1) in RM7. Equation (2) becomes (where $[G] = 1/2$).

$$[P]_{sp} = \frac{k_{17}[M]}{k_2} = \frac{k_{17}k_{15}}{2k_2k_{16}}. \quad (73)$$

Also,

$$K_{GP} = \frac{1}{[P]_{sp}} = \frac{2k_2k_{16}}{k_{15}k_{17}}. \quad (74)$$

This optimized expression for K_{GP} was substituted along with eqn (5) into eqn (23). The resulting $[P]_{max} : G_{min}$ and $[P]_{min} : G_{max}$ pairs were substituted into eqn (72), yielding eqns (75) and

(76) when solved for k_3 .

$$k_3 = \frac{1}{2} \frac{k_{17}k_{15}(2k_4 + 3k_2\varepsilon + 3k_4\varepsilon + k_2\varepsilon^2 + k_4\varepsilon^2)}{k_2k_{16}(\varepsilon + 2)}, \quad (75)$$

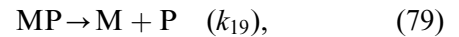
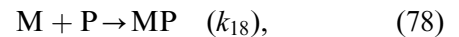
$$k_3 = \frac{1}{2} \frac{k_{17}k_{15}(-2k_4 + 3k_2\varepsilon + 3k_4\varepsilon - k_2\varepsilon^2 - k_4\varepsilon^2)}{k_2k_{16}(\varepsilon - 2)}. \quad (76)$$

These two equations replace eqns (6) and (7) for this mechanism. The k_{3max} , k_{4max} , and A_{unsync} values shown in Table 1 were obtained by setting $k_4 = 0$ in eqn (75) and $k_3 = 0$ in eqn (76). These values are identical to those of RM2 when

$$\frac{k_{17}k_{15}}{k_{16}} = k_1. \quad (77)$$

REGULATORY MECHANISM EIGHT (RM8, "TRANSCRIPTIONAL AND TRANSLATIONAL FEEDBACK")

Regulatory systems employing mRNA can improve regulatory effectiveness by including negative feedback control of translation (Axe & Bailey, 1994; Smolen *et al.*, 2000b). In RM8, P binds M, increasing the rate of M-degradation (Fig. 1). At low $[P]$ (relative to a binding constant for the reaction $P + M \rightleftharpoons MP$), most M's are unbound and active in catalysing the synthesis of more P. At high $[P]$, M's become bound and susceptible to degradation. This decreases the average lifetime of M, and lowers the steady-state concentration of P. The most effective behavior of this system occurs when $k_{16} = 0$, since this will make the MP feedback-control most dramatic. Under this condition, the defining reactions and conservation relationship of RM8 includes eqns (11)–(13), (16)–(18), (65), (67), (78)–(80) and (18).



The corresponding differential equations are eqn (20) and

$$\begin{aligned} d[P]/dt = & k_{17}[M] - k_2[P] + k_3 - k_4[P] - k_5[G][P] \\ & + k_6[GP] - k_{18}[P][M] + k_{19}[MP] + k_{20}[MP], \end{aligned} \quad (81)$$

$$d[M]/dt = k_{15}[G] - k_{18}[P][M] + k_{19}[MP] \quad (82)$$

$$d[MP]/dt = k_{18}[P][M] - k_{19}[MP] - k_{20}[MP] \quad (83)$$

Corresponding steady-state equations include eqns (21), (70) and

$$k_{15}[G] + k_{19}[MP] = k_{18}[P][M], \quad (84)$$

$$k_{18}[P][M] = (k_{19} + k_{20})[MP]. \quad (85)$$

Equations (84) and (85) can be rearranged to eliminate $[MP]$, yielding

$$[M] = \frac{k_{15}[G](k_{20} + k_{19})}{k_{18}k_{20}[P]}. \quad (86)$$

Equation (86) can be substituted into eqn (70), yielding

$$\begin{aligned} [P] &= \frac{k_{17}[M] + k_3}{k_2 + k_4} \\ &= \frac{\sqrt{k_3^2 + 4k_c[G](k_2 + k_4)} + k_3}{2(k_2 + k_4)}, \end{aligned} \quad (87)$$

where

$$k_c = \frac{k_{15}k_{17}(k_{19} + k_{20})}{k_{18}k_{20}}. \quad (88)$$

In the absence of perturbations,

$$[P]_{sp} = \sqrt{\frac{k_c}{2k_2}}. \quad (89)$$

The appropriate equilibrium and $[G]/[G]_{tot}$ relationships are eqns (22) and (23), respectively. Maximum sensitivity occurs at $[P]_{sp}$ when

$$K_{GP} = \frac{1}{[P]_{sp}} = \sqrt{\frac{2k_2}{k_c}}. \quad (90)$$

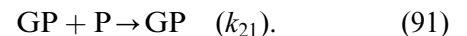
This relationship for K_{GP} was substituted along with eqn (5) into eqn (23), yielding eqn (25). The resulting $[P_{max} : G_{min}]$ and $[P_{min} : G_{max}]$ pairs were substituted into eqn (87), and k_{3max}, k_{4max} and A_{unsync} values (Table 1) were determined as with RM7.

Rate constants were fixed at values that allowed $[M]_{sp} = [MP]_{sp} = 2$ under unperturbed $[P]_{sp}$ conditions. k_{20} was set to the value of k_{16} in RM7 (namely $k_{20} = 1 \text{ min}^{-1}$), allowing k_{15} and k_{17} to be set at RM7 values ($k_{15} = 4 \text{ min}^{-1}$, $k_{17} = 5 \text{ min}^{-1}$). Rate constant k_{19} was arbitrarily set to 1 min^{-1} , resulting in k_{18} with the value 0.002 min^{-1} . Under these conditions (and with $\varepsilon = 0.1$), A_{unsync} is 2.8 times larger than that for RM7 (Table 1).

REGULATORY MECHANISM NINE (RM9, "DEGRADATION NEGATIVE FEEDBACK")

RM2–RM8 probed the effects of regulatory elements that control the rate of P *synthesis*. In contrast, RM9 (Fig. 1) addresses the effect of adding a regulatory element to control the rate of P *degradation*.

RM9 is a simplified form of a situation in which a protease for a target protein P is itself feedback regulated at the transcriptional level, by having P activate the gene encoding the protease (as well as having it repress the promoter for G, its own gene). For optimal regulation, the equilibrium constant describing the binding of P to the protease promoter will equal K_{GP} (and both will equal $1/[P]_{sp}$). Including two genes that interact with P in this manner would afford a rather complicated mechanism in which comparison with other mechanisms would be difficult. However, equivalent behavior can be obtained by letting GP represent the activated P-bound form of the protease promoter, and a catalyst for the synthesis of the protease. The concentration of the protease will be proportional to $[GP]$. The defining reactions, rate constants, and conservation relationship for RM9 include eqns (10), (12), (13), (16)–(18) and



The corresponding differential equations include eqn (20) and

$$\begin{aligned} d[P]/dt = & k_1[G] - k_{21}[P][GP] + k_3 - k_4[P] \\ & - k_5[G][P] + k_6[GP]. \end{aligned} \quad (92)$$

Steady-state, equilibrium, and $[G]/[G]_{tot}$ relationships include eqns (21)–(23) and (93).

$$k_1[G] + k_3 = k_{21}[P][GP] + k_4[P]. \quad (93)$$

Combining eqns (18) and (93) to eliminate $[GP]$ yields

$$[P] = \frac{k_1[G] + k_3}{k_{21}[GP] + k_4} = \frac{k_1[G] + k_3}{k_{21} - k_{21}[G] + k_4}. \quad (94)$$

This equation replaces eqn (1). In the absence of perturbations and at $[G] = 1/2$,

$$[P]_{sp} = \frac{k_1}{k_{21}} = \frac{1}{K_{GP}}. \quad (95)$$

This relationship for K_{GP} and eqn (5) were substituted into eqn (23) yielding the G_{min} and G_{max} values in eqn (25). $[P_{max} : G_{min}]$ and $[P_{min} : G_{max}]$ pairs were substituted into eqn (94) and the resulting equations were solved for k_{3max} , k_{4max} , and A_{unsync} as in other mechanisms. To keep $[P]_{sp} = 1000$ and $k_1 = 20 \text{ min}^{-1}$, k_{21} was set to 0.02 min^{-1} . Under these constraints, A_{unsync} (Table 1) is 1.8 times greater than that for RM2.

Discussion

REGULATORY MECHANISMS LACKING NEGATIVE FEEDBACK

Systems consisting of synthesis and degradation processes without negative feedback relationships (e.g. RM1) regulate $[P]$ to some extent. Steady-state regulatory effectiveness, as reflected by the area of the unsynchronized regulated region of $k_3 : k_4$ perturbation space, was proportional to the biosynthesis and degradation rates of P (i.e. $A_{unsync} \propto k_1 \times k_2$). Thus, most effective regulation is achieved with *rapid* rates of synthesis and degradation. In real cells, many proteins are stable towards proteolysis, and so the rate of “degradation” essentially corresponds to the rate of dilution that accompanies cell growth (Axe & Bailey, 1994). This situation

limits the degree to which such proteins can be regulated under steady-state conditions; i.e. with everything else equal, stable proteins are not regulated as well as rapidly degraded ones.

The ratio of synthesis rates to degradation rates (i.e. k_1/k_2) determines the steady-state “set-point” copy number (or concentration) in cells (Hargrove, 1994). Proteins present at high copy number are either produced faster or degraded more slowly than those present in lesser amounts. The *relative* effectiveness by which a protein is regulated is not related to copy number *per se*.

EFFECT OF CHANGING THE REGULATORY TIGHTNESS PARAMETER ε

The magnitude of ε has a significant influence on the size of A_{unsync} . In RM1, if ε is “relaxed” from 0.1 to 0.3, A_{unsync} increases 11.6 times! The effect is qualitatively similar for the other mechanisms. Thus, there is an advantage to setting ε as high as can be tolerated by the application. For example, if P has a first-order rate-law dependence for all reactions for which it is involved, a $\pm 30\%$ deviation may be acceptable, as this would translate into “only” a $\pm 30\%$ deviation in rate. However, if P exhibits higher-order dependencies, the same deviation in $[P]$ would cause disproportionately greater deviations in rate. In such cases, smaller ε values would be required.

NEGATIVE FEEDBACK, DIMERIZATION, COOPERATIVITY, AND BOOLEAN MECHANISMS

We evaluated the effect of negative feedback relationships occurring in transcription, translation, and degradation processes. For transcription, we examined basic negative feedback (RM2), multiple sequential binding feedback (RM3), Boolean feedback (RM4), dimerization feedback (RM5), and a combination of multiple sequential binding and dimerization feedback (RM6). Regulatory effectiveness was related to the slope of the $[G]/[G]_{tot}$ vs. $\log[P]$ plots at $[P]_{sp}$ and $[G]/[G]_{tot} = 1/2$ (Fig. 2). RM4 with Boolean feedback (infinite slope) was most effective in regulating $[P]$, while that without feedback (RM1, with slope=0) was least effective. The system with two P’s binding G sequentially was

as effective as the one where two P's dimerize and the dimer binds G (as long as the second P binds GP very tightly). Under these conditions both were about twice as effective as the mechanism involving basic negative feedback, consistent with their corresponding slopes being twice as great. The system with both multiple sequential binding and dimerization was about twice as effective as those with either regulatory element alone. Additional improvements would be expected by including further oligomerization steps (e.g. the binding of two dimers to form a tetramer). Real regulatory systems often exhibit both multiple binding and oligomerization of transcription factors (John *et al.*, 1999; Payre *et al.*, 1997). Feedback terms that were added to either translation or degradation processes also improved regulatory effectiveness. Although our analysis does not indicate the exact effect of combining feedback terms, each term that was added increased regulatory effectiveness roughly by a factor of two. A more sophisticated analysis is required to determine whether this observation can be used to assess the regulatory abilities of mechanisms with other combinations of regulatory elements.

TRANSCRIPTION/TRANSLATION CASCADES

The regulatory effectiveness of mechanisms that include transcription/translation cascades may naively appear no better than those lacking them (see Table 1), but this results from setting rate constants according to eqn (77). In hypothetical cascades, the rate of E_2 may be adjusted to any value (including the amplified rate). In contrast, the rates of real systems are constrained by physical limitations. The assigned value of k_1 (i.e. 20 min^{-1}) is calculated from estimates of actual rates of protein biosynthesis. However, these rates include amplification due to the transcription/translation cascades present in real cells. In our mechanisms, the amplification factor equaled the steady-state concentration of M. Typically there are ~ 2 copies of an mRNA per cell (Neidhardt & Umbarger, 1996). If there were $1/2$ free G per cell, eqn (71) indicates that $k_{15}/k_{16} \sim 4$. Then, eqn (77) requires that $k_{17} \sim 5 \text{ min}^{-1}$. Since the half-life of mRNA can be as short as $\sim 1/2 \text{ min}^{-1}$ (Wagner, 2000;

Kushner, 1996; Arraiano *et al.*, 1988; Alberts *et al.*, 1994), we assume $k_{16} \sim 1 \text{ min}^{-1}$, yielding $k_{15} \sim 4 \text{ min}^{-1}$. The homeostatic ability of these mechanisms, as defined by A_{unsync} , is proportional to $[M]$, so the effect of this cascade (with $[M] \sim 2$) is to *double* regulatory effectiveness.

The improvement in regulatory effectiveness resulting from transcription/translation cascades is probably even greater in real systems because they contain additional levels. Many ribosomes typically bind simultaneously to the same mRNA strand, giving rise to polysomes. Since each ribosome:RNA complex is a catalyst for protein synthesis, the presence of polysomes adds another cascade level, in this case amplified by the number of such complexes per mRNA strand. Also, in many real systems, transcription factors other than P itself bind the locus responsible for expressing the P-encoding gene. These factors typically bind a small molecule (Q) whose concentration is influenced by P (for example, Q may be the product of a P-catalysed reaction). This arrangement would constitute another level of the cascade structure of transcription/translation processes. Feedback/cooperativity relationships may then be added at each level. Ferrell, Kholodenko, and others have shown that including such elements affords switch-like response-curves (Ferrell, 1996–1998; Kholodenko *et al.*, 1997; Kholodenko, 2000; Huang & Ferrell, 1996; Brown *et al.*, 1997).

SUMMARY

The results of our analysis reveal some principles for constructing effective homeostatic regulatory mechanisms for a generic protein P operating under steady-state conditions. They include:

- Use rapid rates of P synthesis and degradation. Unstable proteins can be more effectively regulated than stable ones. However, there is a trade-off between regulation and energy costs; better regulation is costly. Consequently, many proteins in cells are stable and not regulated as well as they would be if they were unstable.
- Use a negative-feedback-controlled cascade structure, in which the ultimate product of the cascade (P in this case) is also the signaling molecule that inhibits the first-level catalyst (G

in this case). This structure is commonly applied to the synthesis “side” of regulatory mechanisms. However, if P (or a transcription factor sensitive to [P]) activates a gene that encodes a protease that degrades P, a cascade structure is also present on the degradation “side” of the mechanism. These cascade structures can amplify the rates of synthesis (and degradation), thereby improving regulation.

- Apply feedback relationships wherever possible, including at each level of each cascade. In these relationships, P (or a related transcription factor) either binds or unbinds the catalyst operating at each level. The equilibrium binding constants should be matched such that half of each type of catalyst is bound at the desired set-point concentration of P.

- Allow multiple copies of P (or a related transcription factor) to bind sequentially. This adds cooperativity and increases the sensitivity of the feedback, such that smaller percentage changes in [P] cause greater percent changes in the relative concentration of the active form of the catalyst (e.g. $[G]/[G]_{tot}$). Another way of adding cooperativity and ultrasensitivity is by allowing P (or transcription factors) to oligomerize. Higher-level oligomers yield more effective regulation (i.e. mechanisms using tetramers regulate better than those using dimers). This improvement in regulation explains why most if not all transcription factors bind as oligomers.

LINKING REGULATORY MODULES

Regulatory networks in cells are poorly understood (Smolen *et al.*, 2000a,b). They appear to be organized hierarchically, from genes at the lowest level to operons, regulons, and modulons at higher levels (Neidhardt & Savageau, 1996), and have been viewed analogously to electronic circuits (McAdams & Arkin, 1998; McAdams & Shapiro, 1995). Whether regulatory networks are best viewed as consisting of exceedingly large numbers of strongly interacting components or of smaller numbers of weakly interacting groups or *modules* is unknown, but the latter possibility seems more computationally tractable. Bray (1995) has discussed the possibility of linking small sets of

protein-based reactions to perform logical computations. Thieffry & Romero (1999) found that as the size of a Boolean circuit increases, the fraction of parameter combinations that lead to functionality declines, suggesting that networks are composed of loosely interacting modules. Others have proposed to view regulatory networks as hierarchically organized functional modules that can be analyzed individually and combined (Lengeler, 2000; Kremling *et al.*, 2000; Hofmeyr & Westerhoff, 2001).

Our analysis suggests a method of linking regulatory modules in the design of complex networks. Consider two modules that regulate P_1 and P_2 within a small deviation of set-point concentrations. To link these modules, P_1 and P_2 must interact. For example, P_1 and P_2 may bind each other, or P_1 may catalyze a reaction involving P_2 . In any event, the effect of the interaction would be to alter the concentration of one or both P's. Importantly, each module in the link would view the effect of the other as a perturbation. Thus, as long as the linking reaction was slow relative to the core rates by which P_1 and P_2 were regulated (i.e. keeping both systems within their unsynchronized perturbation-space rectangles), the two P's in the linked system would remain regulated. Additional modules could be linked under similar restrictions, eventually forming a regulatory network consisting of semi-autonomous regulatory modules with rapid intra-module reactions and slow inter-module or linking reactions (“strength” of interactions within and between modules would be defined kinetically). Assuming random interactions between modules, we suspect that the strength of the interactions would need to diminish as the number of linked modules increases. However, multiple perturbations would combine randomly (some increasing P's and some decreasing them). To *assure* that individual perturbations would not combine to yield an overall perturbation that exceeds the regulatory capacity of any individual module, the magnitude of each linking interaction would have to decline as the number of linkage reactions increased. However, the probability that random perturbations would combine exclusively in one direction or the other would also decrease as the number of interactions

increase, such that it might be possible to relax this restriction.

Whether this “weak-coupling-of-regulatory-modules” approach could successfully mimic global regulatory networks in cells is unknown. If the strength (i.e. rate) of the inter-module linking reactions were as strong (i.e. fast) as the intra-module core reactions, the rate constants used in constructing individual modules would have little bearing on those required in the network. In this case, it would appear impossible to randomly select the exceedingly large number of rate constants and copy numbers required to elicit the observed regulatory behavior (assuming the correct model). Most if not all such data would need to be measured experimentally, an arduous if not impossible task. On the other hand, if global regulatory networks consisted of weakly linked semi-autonomous regulatory modules, it might be possible to model the behavior of the network by first modeling the behavior of individual unlinked modules. The relatively small number of rate constants associated with individual modules could be estimated from the ideal or assumed behavior (for example, homeostasis), essentially ignoring perturbing interactions. The kinetic parameters (rate constants and copy numbers) required to mimic the behavior of unlinked modules would be retained after linking. Once hypothetical networks were assembled and “working”, behavior could be improved by randomly perturbing these parameters in the manner described by Bray (1995). Even if these hypothetical regulatory modules operated by different mechanisms *vis-à-vis* actual modules, they could serve as “surrogates” that function similarly (e.g. both regulate P within some degree of effectiveness). Once the actual mechanism and/or kinetic data became available, surrogates could be replaced without disrupting the entire network. Given the complexities involved, such a piecemeal approach may represent one of the few viable ways of modeling the behavior of global regulatory networks in living systems.

We thank Deborah Siegle for critically reading this manuscript prior to publication and for helpful discussion. The National Institutes of Health (GM63958), Pfizer Pharmaceuticals Inc (fellowship

to C.S.), and the National Science Foundation (Grant 008336) sponsored this study. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsors.

REFERENCES

- ALBERTS, B., BRAY, D., LEWIS, J., RAFF, M., ROBERTS, K. & WATSON, J. D. (1994). *Molecular Biology of the Cell*. 3rd Edn., p. 234, New York: Garland Publishing, Inc.
- ARRAIANO, C. M., YANCEY, S. D. & KUSHNER, S. R. (1988). Stabilization of discrete mRNA breakdown products in *ams pnp rnb* multiple mutants of *Escherichia coli* K-12. *J. Bacteriol.* **170**, 4725–4633.
- AXE, D. D. & BAILEY, J. E. (1994). Modeling the regulation of bacterial genes producing proteins that strongly influence growth. *Biotech. Bioeng.* **43**, 242–257.
- BARKAI, N. & LEIBLER, S. (1997). Robustness in simple biochemical networks. *Nature* **387**, 913–917.
- BECSKEI, A. & SERRANO, L. (2000). Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593.
- BOWER, J. M. & BOLOURI, H. (ed.) (2001). *Computational Modeling of Genetic and Biochemical Networks*. Cambridge MA: MIT Press.
- BRAY, D. (1995). Protein molecules as computational elements in living cells. *Nature* **376**, 307–312.
- BROWN, G. C., HOEK, J. B. & KHOLODENKO, B. N. (1997). Why do protein kinase cascades have more than one level? *Trends Biochem. Sci.* **22**, 288.
- CHEN, T., HE, H. L. & CHURCH, G. M. (1999). Modeling gene expression with differential equations. *Pacific Symposium of Biocomputing*, Hawaii, pp. 29–40.
- CHOCK, P. B., RHEE, S. G. & STADTMAN, E. R. (1980). Interconvertible enzyme cascades in cellular regulation. *Ann. Rev. Biochem.* **49**, 813–843.
- EDWARDS, R. & GLASS, L. (2000). Combinatorial explosion in model gene networks. *Chaos* **10**, 691–704.
- FERRELL Jr. J. E. (1996). Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends Biochem. Sci.* **21**, 460–466.
- FERRELL Jr. J. E. (1997). How responses get more switch-like as you move down a protein kinase cascade. *Trends Biochem. Sci.* **22**, 288–289.
- FERRELL Jr. J. E. (1998). How regulated protein translocation can produce switch-like responses. *Trends Biochem. Sci.* **23**, 461–465.
- GLASS, L. (1975). Classification of biological networks by their qualitative dynamics. *J. theor. Biol.* **54**, 85–107.
- GLASS, L. & KAUFMAN, S. A. (1973). The logical analysis of continuous non-linear biochemical control networks. *J. theor. Biol.* **39**, 103–129.
- GOLDBETER, A. & KOSHLAND Jr. D. E. (1982). Sensitivity amplification in biochemical systems. *Q. Rev. Biophys.* **15**, 555–591.
- GOTTESMAN, S. & MAURIZI, M. R. (1992). Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.* **56**, 592–621.
- HARGROVE, J. L. (1994). *Kinetic Modeling of Gene Expression*. Austin, TX: R.G. Landes Co.
- HLAVACEK, W. S. & SAVAGEAU, M. A. (1996). Rules for coupled expression of regulator and effector genes in inducible circuits. *J. Mol. Biol.* **225**, 121–139.

- HOFMEYR, J. H. S. & CORNISH-BOWDEN, A. (1996). Co-response analysis: A new experimental strategy for metabolic control analysis. *J. theor. Biol.* **182**, 371–380.
- HOFMEYR, J. H. S. & WESTERHOFF, H. V. (2001). Building the cellular puzzle—control in multilevel reaction networks. *J. theor. Biol.* **208**, 261–285.
- HUANG, C.-Y. F. & FERRELL Jr. J. E. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl Acad. Sci. U.S.A.* **93**, 10078–10083.
- JOHN, S., VINKEMEIER, U., SOLDAINI, E., DARNELL Jr. J. E. & LEONARD, W. J. (1999). The significance of tetramerization in promoter recruitment by Stat5. *Mol. Cell. Biol.* **19**, 1910–1918.
- KHOLODENKO, B. N. (2000). Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur. J. Biochem.* **267**, 1583–1588.
- KHOLODENKO, B. N., HOEK, J. B., WESTERHOFF, H. V. & BROWN, G. C. (1997). Quantification of information transfer via cellular signal transduction pathways. *FEBS Lett.* **414**, 430–434.
- KOSHLAND Jr. D. E., GOLDBETER, A. & STOCK, J. B. (1982). Amplification and adaptation in regulatory and sensory systems. *Science* **217**, 220–225.
- KREMLING, A., JAHREIS, K., LENGELER, J. W. & GILLES, E. D. (2000). The organization of metabolic reaction networks: a signal oriented approach to cellular models. *Metabolic Eng.* **2**, 1–11.
- KUSHNER, S. R. (1996). mRNA decay. In: Neidhardt, F. C., (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd Edn, Vol. 1, Washington DC: ASM Press, Chapter 56.
- LAUB, M. T., MCADAMS, H. H., FELDBLYUM, T. & SHAPIRO, L. (2000). Global analysis of the genetic network controlling a bacterial cell cycle. *Science* **290**, 2144–2148.
- LENGELER, J. W. (2000). Metabolic networks: a signal-oriented approach to cellular models. *Biol. Chem.* **381**, 911–920.
- LEWIS, B. (1998). *Genes VI*. New York: Oxford University Press.
- MCADAMS, H. H. & ARKIN, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. U.S.A.* **94**, 814–819.
- MCADAMS, H. H. & ARKIN, A. (1998). Simulation of prokaryotic genetic circuits. *Ann. Rev. Biophys. Biomol. Struct.* **27**, 199–224.
- MCADAMS, H. H. & ARKIN, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics* **15**, 65–69.
- MCADAMS, H. H. & SHAPIRO, L. (1995). Circuit simulation of genetic networks. *Science* **269**, 650–656.
- MESTL, T., PLAhte, E. & OMHOLT, S. W. (1995). A mathematical framework for describing and analyzing gene regulatory networks. *J. theor. Biol.* **176**, 291–300.
- MEISKE, W. & REICH, J. G. (1987). Mathematical description of the control of steady-states in metabolic systems. *Syst. Anal. Model. Simulation* **4**, 423–435.
- NEIDHARDT, F. C. & SAVAGEAU, M. A. (1996). Regulation beyond the operon. in: Neidhardt, F. C., (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, Vol. 1, 2nd Edition, Washington, DC: ASM Press, Chapter 84.
- NEIDHARDT, F. C. & UMBARGER, H. E. (1996). Chemical composition of *E. coli*. In: Neidhardt, F. C., (Ed.), *Scherichia Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd Edn, Vol. 1, Washington, DC: ASM Press, Chapter 3.
- NI, T.-C. & SAVAGEAU, M. A. (1996). Model assessment and refinement using strategies from biochemical systems theory: application to metabolism in human red blood cells. *J. theor. Biol.* **179**, 329–368.
- OMHOLT, S. W., KEFANG, X., ANDERSEN, Ø. & PLAhte, E. (1998). Description and analysis of switchlike regulatory networks exemplified by a model of cellular iron homeostasis. *J. theor. Biol.* **195**, 339–350.
- PAYRE, F., BUONO, P., VANZO, N. & VINCENT, A. (1997). Two types of zinc fingers are required for dimerization of the serendipity δ transcriptional activator. *Mol. Cell. Biol.* **17**, 3137–3145.
- PTASHNE, M. (1992). *A Genetic Switch: Phage λ and Higher Organisms*, 2nd Edn., Oxford: Cell Press & Blackwell Scientific Publications.
- ROSEN, R. (1968). Recent developments in the theory of control and regulation of cellular processes. In: Bourne, G. H., Danielli, J. F., (eds.), *International Review of Cytology*. pp. 25–88, New York: Academic press.
- SAVAGEAU, M. A. (1998). Demand theory of gene regulation. I. Quantitative development of the theory. *Genetics* **149**, 1665–1676.
- SHACTER, E., BOON CHOCK, P. & STADTMAN, E. R. (1984). Regulation through phosphorylation/dephosphorylation cascade systems. *J. Biol. Chem.* **259**, 12252–12259.
- SHACTER-NOIMAN, E., CHOCK, P. B. & STADTMAN, E. R. (1983). Protein phosphorylation as a regulatory device. *Philos Trans. R. Soc. London B* **302**, 157–166.
- SMOLEN, P., BAXTER, D. A. & BYRNE, J. H. (2000a). Mathematical modeling of gene networks. *Neuron* **26**, 567–580.
- SMOLEN, P., BAXTER, D. A. & BYRNE, J. H. (2000b). Modeling transcriptional control in gene networks—methods, recent results, and future directions. *Bull. Math. Biol.* **62**, 247–292.
- THIEFFRY, D. & ROMERO, D. (1999). The modularity of biological regulatory networks. *Biosystems* **50**, 49–59.
- THOMAS, R. (1973). Boolean formalization of genetic control circuits. *J. theor. Biol.* **42**, 563–585.
- THOMAS, R. (1991). Regulatory networks as asynchronous automata: a logical description. *J. theor. Biol.* **153**, 1–23.
- THOMAS, R., THIEFFRY, D. & KAUFMAN, M. (1995). Dynamical behavior of biological regulatory networks—I. Biological role of the feedback loops and practical use of the concept of the loop-characteristic state. *Bull. Math. Biol.* **57**, 247–276.
- TYSON, J. & OTHMER, H. G. (1978). The dynamics of feedback control circuits in biochemical pathways. *Prog. theor. Biol.* **5**, 2–62.
- VOET, D. & VOET, J. G. (1995). *Biochemistry*, 2nd Edn., p. 1010. New York: John Wiley & Sons, Inc.
- WAGNER, R. (2000). *Transcription Regulation in Prokaryotes*, p. 91. New York: Oxford University Press.
- WOLF, D. M. & EECKMAN, F. H. (1998). On the relationship between genomic regulatory element organization and gene regulatory dynamics. *J. theor. Biol.* **195**, 167–186.