Supporting Text

Derivation of the Governing Model Equations. The chemical reactions describing the autocatalytic network are naturally divided into two categories: fast and slow. The fast reactions have rate constants of order seconds and are therefore assumed to be in equilibrium with respect to the slowly changing variables, which evolve on time scales of order minutes. If we let X, X_2, and D_i denote the repressor, repressor dimer, and DNA promoter sites with i dimers bound, then we may write the fast reactions as

$$\begin{aligned} X + X &\underset{k_{-1}}{\overset{k_1/V}{\rightleftharpoons}} X_2 \\ D_0 + X_2 &\underset{k_{-2}}{\overset{k_2/V}{\rightleftharpoons}} D_1 \\ D_1 + X_2 &\underset{k_{-3}}{\overset{k_3/V}{\rightleftharpoons}} D_2 \\ D_2 + X_2 &\underset{k_{-4}}{\overset{k_4/V}{\rightleftharpoons}} D_3 \end{aligned} \qquad [1]$$

where the volume dependence is explicit in the reaction rates because the volume is a slowly evolving function of time due to the growth of the host cell.

The slow irreversible reactions are transcription, translation, and protein denaturation or destabilization. If no repressor is bound to the operator region or if a single repressor dimer is bound to the first right operator site (OR1), transcription proceeds at a basal rate. If, however, a repressor dimer is bound to OR2, the binding affinity of RNA polymerase to the promoter region is enhanced by a factor α, leading to an amplification of transcription. Denaturation is due to the temperature-induced

destabilization of CI monomers. We write the reactions governing these processes as

$$\begin{aligned} D_0 &\xrightarrow{k_t} D_0 + X \\ D_1 &\xrightarrow{k_t} D_1 + X \\ D_2 &\xrightarrow{\alpha k_t} D_2 + X \\ X &\xrightarrow{k_d} \quad , \end{aligned} \qquad [2]$$

where transcription and translation are modeled as a single reaction with rate constant k_t.

The cellular volume increase and division are modeled as follows. For time just after cell division to just before, we let the volume increase as

$$V = V_0 \; e^{\ln(2) \; T/\tau_0}, \qquad [3]$$

where V_0 is the volume of the host cell at the beginning of its growth phase, and τ_0 is the cell-division time. At times $T = q\tau_0$, where q is an integer, we let $V \to V/2$ and $n \to n/2$ for each of the protein species. This operation represents the halving of the volume at division along with the redistribution of molecules to one daughter cell. Defining the dimensionless variables $v = V/V_0$ and $t = T/\tau_0$, we have

$$v = e^{\ln(2) \; t} \qquad [4]$$

so that, in these units, time is measured in terms of the cell-division time ($t \; \epsilon \; [0, 1]$), and the volume oscillates between 1 and 2.

We seek to write a dynamical equation for the slowest process described by the reactions in Eqs. 1 and 2. The association and dissociation rates for dimerization and protein–DNA binding are fast compared with transcription and degradation. Thus, although cI monomers are transcribed and degraded at a slow rate, they are not a slowly changing quantity because they quickly equilibrate with cI dimers (1). The slowly changing variable is the total number of cI molecules (in any form), and is defined as

$$z = x + 2x_2 + 2d_1 + 4d_2 + 6d_3, \qquad [5]$$

where x, x_2, and d_i are the numbers of cI monomers, dimers, and dimer–DNA complexes. The temporal evolution of the total number of cI molecules is given by the rate equation

$$\frac{dz}{dt} = \beta(d_0 + d_1) + \alpha\beta d_2 - \gamma_x x, \qquad [6]$$

where $\beta \equiv k_t\tau_0$ and $\gamma_x \equiv k_d\tau_0$ are dimensionless parameters representing the number of cI monomers produced (basally) and destabilized over the course of one cell-division time.

With respect to the slowly evolving process described by Eq. 6, the fast reactions given by Eq. 1 can be considered to be in equilibrium. This gives the following algebraic relations,

$$\begin{aligned} x_2 &= \frac{k_1}{k_{-1}V}x^2 \qquad [7] \\ d_1 &= \frac{k_2}{k_{-2}V}d_0x_2 \\ d_2 &= \frac{k_3}{k_{-3}V}d_1x_2 \\ d_3 &= \frac{k_4}{k_{-4}V}d_2x_2. \end{aligned}$$

Defining dimensionless equilibrium constants* as $c_i \equiv k_i/(k_{-i}V_0A)$, where A is the Avogadro number, and letting $c \equiv c_1c_2$, we have

$$\begin{aligned} x_2 &= c_1x^2/v \qquad [8] \\ d_1 &= cd_0x^2/v^2 \\ d_2 &= \sigma_1c^2d_0x^4/v^4 \\ d_3 &= \sigma_1\sigma_2c^3d_0x^6/v^6, \end{aligned}$$

where the σ_i prefactors denote the relative affinities for dimer binding to OR1 versus that of binding to OR2 (σ_1) and OR3 (σ_2).

We can eliminate the number of unbound operator sites d_0 by taking the copy-number m as constant,

$$m = d_0 + d_1 + d_2 + d_3. \qquad [9]$$

This leads to an equation for d_0:

$$d_0(x,v) = \frac{m}{1 + cx^2/v^2 + \sigma_1 c^2 x^4/v^4 + \sigma_1\sigma_2 c^3 x^6/v^6}. \quad [10]$$

Now, utilizing Eqs. 8 and 10, we can rewrite Eq. 6 as

$$\dot{z} = \frac{m\beta(1 + cx^2/v^2 + \alpha\sigma_1 c^2 x^4/v^4)}{1 + cx^2/v^2 + \sigma_1 c^2 x^4/v^4 + \sigma_1\sigma_2 c^3 x^6/v^6} - \gamma_x x. \quad [11]$$

The last step is to write $\dot{z}$ in terms of $\dot{x}$. Utilizing Eqs. 5 and 8, we obtain

$$\begin{aligned} \dot{z} &= \dot{x}\left[1 + \frac{4c_1 x}{v} + \frac{4cxd_0(x,v)}{v^2} + \frac{16\sigma_1 c^2 x^3 d_0(x,v)}{v^4} + \frac{36\sigma_1\sigma_2 c^3 x^5 d_0(x,v)}{v^6}\right] \quad [12] \\ &+ \dot{d}_0\left(2\frac{cx^2}{v^2} + 4\frac{\sigma_1 c^2 x^4}{v^4} + 6\frac{\sigma_1\sigma_2 c^3 x^6}{v^6}\right), \end{aligned}$$

where we have assumed that the volume increases adiabatically so that $\dot{v}$ terms can be ignored with respect to the time derivative of x. Additionally, provided the number of monomers is not too small[†], the number of operator states d_0 without bound dimers will be small and rarely changing. We thus can approximate Eq. 12 by letting $\dot{d}_0(x,v) = 0$. This leads to the governing equation for the evolution of the number of cI monomers,

$$\dot{x} = \frac{1}{h(x,v)}(\beta f(x,v) - \gamma_x\, x), \quad [13]$$

where

$$f(x,v) = \frac{m(1 + cx^2/v^2 + \alpha\sigma_1 c^2 x^4/v^4)}{1 + cx^2/v^2 + \sigma_1 c^2 x^4/v^4 + \sigma_1\sigma_2 c^3 x^6/v^6} \quad [14]$$

and

$$h(x,v) = 1 + \frac{4c_1 x}{v} + \frac{4cxd_0(x,v)}{v^2} + \frac{16\sigma_1 c^2 x^3 d_0(x,v)}{v^4} + \frac{36\sigma_1\sigma_2 c^3 x^5 d_0(x,v)}{v^6}. \quad [15]$$

The equation describing the evolution of the number of GFP molecules can be deduced by noting that, because GFP proteins are translated in tandem with cI proteins, the governing equation is analogous to Eq. 6,

$$\frac{dg}{dt} = \eta\beta(d_0 + d_1) + \alpha\beta d_2 - \gamma_g g, \quad [16]$$

where the parameter η is the relative efficiency of production of GFP compared with that of cI. The reduction of this equation follows the same route as that described above and leads to

$$\dot{g} = \eta\beta f(x,v) - \gamma_g\, g. \quad [17]$$

Internal Fluctuations. Internal fluctuations arise from the small number of reactant molecules. Within the context of gene regulation, one recently employed technique for incorporating such fluctuations is through the direct numerical simulation of the reactions given by Eqs. 1 and 2 (2,3). In the present work we adopt an alternative approach. This approach amounts to the generalization of the deterministic equations to include stochastic terms describing the internal fluctuations. Within this framework, Eqs. 13 and 17 describe the evolution of the mean number of cI and GFP proteins, and the stochastic terms can be viewed as perturbations away from the steady states characterized by the mean values. The result is a generalization of Eqs. 13 and 17 to the following Langevin equations,

$$\begin{aligned}
\dot{x} &= \frac{1}{h(x,v)}\{\beta f(x,v) - \gamma_x\ x\} + \sqrt{\frac{1}{h(x,v)}\{\beta f(x,v) + \gamma_x\ x\}}\ \xi_x(t) \\
\dot{g} &= \eta\beta f(x,v) - \gamma_g\ g + \sqrt{\eta\beta f(x,v) + \gamma_g\ g}\ \xi_g(t),
\end{aligned}$$

where the $\xi_i(t)$ are rapidly fluctuating random terms with zero mean ($< \xi_i(t) >= 0$), and the statistics of the $\xi_i(t)$ are such that $< \xi_i(t)\xi_j(t') >= \delta_{i,j}(t - t')$.

Our simulation of these equations follows the recipe given in ref. 4. Generally, consider a set of M chemical species, with $n_i(t)$ denoting the number of molecules of species i at time t. Setting $\mathbf{n} = \left[\begin{array}{cccc} n_1 & n_2 & \cdots & n_M \end{array}\right]$, the evolution equation may be written in the form

$$\begin{aligned}
\dot{\mathbf{n}} &= F(\mathbf{n}) - G(\mathbf{n}) + \sqrt{F(\mathbf{n}) + G(\mathbf{n})}\xi(t) && [18] \\
&= P(\mathbf{n}) + Q(\mathbf{n})\xi(t)\,, && [19]
\end{aligned}$$

where $F(\mathbf{n}) = \left[\begin{array}{cccc} F_1(\mathbf{n}) & F_2(\mathbf{n}) & \cdots & F_M(\mathbf{n}) \end{array}\right]$ [similarly for $G(\mathbf{n})$, $P(\mathbf{n})$, and $Q(\mathbf{n})$], $P(\mathbf{n}) \equiv F(\mathbf{n}) - G(\mathbf{n})$, and $Q(\mathbf{n}) \equiv \sqrt{F(\mathbf{n}) + G(\mathbf{n})}$.

For an individual species we have

$$\dot{n}_i = P_i(\mathbf{n}) + Q_i(\mathbf{n})\xi(t). \qquad [20]$$

Updates are computed by using

$$n_i(t+\Delta t) = n_i(t) + P_i(\mathbf{n})\Delta t + Q_i(\mathbf{n})\Xi(t) + \frac{1}{2}[\Xi(t)]^2 \sum_{j=1}^{M} Q_j(\mathbf{n}) \frac{\partial Q_i}{\partial n_j}(\mathbf{n}) \,, \qquad [21]$$

where $\Xi(t)$ is a Gaussian random variable with zero mean and standard deviation $\sigma^2 = \sqrt{\Delta t}$, and $\mathbf{n} = \mathbf{n}(t)$.

GFP Fluorescence and Molecule Number. To directly compare the modeling results with the experiment, we needed to convert the GFP molecule number g obtained from Eq. 18 to a corresponding fluorescence value F. It is known that GFP fluorescence is temperature-dependent (4); therefore we assumed a destabilization-dependent proportionality between the number of GFP molecules and the corresponding fluorescence,

$$F = c(\gamma_x)\ (g + b_0), \qquad [22]$$

where b_0 is a fixed constant and $c(\gamma_x)$ is chosen for each destabilization value.

Results

Simulation results are obtained by numerically integrating Eq. 18 and using Eq. 22 to transform to fluorescence values. Because the experiment makes use of the $\mathrm{P_{RM}}$ promoter region of λ phage, many of the parameters are known from the literature, and in Table 1 we give an extensive list of all parameters used in the simulations. In Fig. 2, we plot the simulated distributions for the GFP fluorescence and cI monomer number alongside of those obtained from the experiment for the entire destabilization (model) and temperature (experiment) sweep.

It is important to note that only two parameters are adjusted in generating the agreement between model and experimental distributions over the entire temperature range (Fig. 2). These adjustable parameters are the destabilization γ_x and the fluorescence proportionality constant $c(\gamma_x)$. As noted in the main text of the article, the fits allow the deduction of the destabilization as a function of temperature, and an exponential fit of these data is in excellent agreement with results reported elsewhere (5). Along these same lines, the results can be taken as a prediction for

the temperature dependence of GFP fluorescence. In Fig. 7 we plot the fluorescence proportionality constant $c(\gamma)$ as a function of both the destabilization and temperature, respectively. Although future experiments are needed to test this prediction, it is interesting to note that the relative values of the proportionality constant for temperatures of 37 and 42°C is $c(37°\mathrm{C})/c(42°\mathrm{C}) = 2.5$, which is in agreement with the value of 2.78 reported for the GFP mutant GFP-A (4). This effect is observed in the control experiments containing the wild–type *cI* gene, presented in Fig. 1D. In additional control experiments, the maturation rate of GFP was observed to occur on the order of tens of minutes, consistent with previously reported studies (6,7).

* Since the number of proteins is typically of order 10^3 or less, these equilibrium constants are the natural units.

† This assumption is consistent with the validity of the rate equation approach. Indeed if the number of monomers is too small, then the rate equation approach is not valid, and techniques originating from the general master equation must be employed (see ref. 1).

1. Kepler, T. & Elston, T. (2001) *Biophys. J.* **81,** 3116–3136.

2. Gillespie, D. (1977) *J. Phys. Chem.* **81,** 2340–2361.

3. Arkin, A., Ross, J. & McAdams, H. H. (1998) *Genetics* **149,** 1633–1648.

4. Siemering, K. R., Golbik, R., Sever, R. & Haseloff, J. (1996) *Curr. Biol.* **6,** 1653–1663.

5. Villaverde, A., Benito, A., Viaplana, E. & Cubarsi, R. (1993) *Appl. Environ. Microbiol.* **59,** 3485–3487.

6. Cluzel, P., Suretter, M. & Leibler, S. (2000) *Science* **287,** 1652–1655.

7. Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. (1999) *Nat. Biotechnol.* **17,** 691–695.