Supporting Information

Schultz *et al.* 10.1073/pnas.0810366105

SI Text

Stochastic Formulation. A more detailed description of the system than that provided by macroscopic kinetics can be obtained by describing all of the possible chemical reactions in the system, including binding and unbinding, as stochastic events. Denoting the unbound and bound promoters of gene a as $\mathcal{A}_1$ and $\mathcal{A}_0$, the binding and unbinding rates as h and f ($X^{eq} = f/h$), xa and x_b being the degree of oligomerization (cooperativity) to which transcription factors bind the DNA, we have the following reactions:

$$\mathcal{A}_1 + x_a b \underset{f_a}{\overset{h_a(b)}{\rightleftharpoons}} \mathcal{A}_0 \qquad \mathcal{B}_1 + x_b a \underset{f_b}{\overset{h_b(a)}{\rightleftharpoons}} \mathcal{B}_0$$

$$\mathcal{A}_1 \overset{g_{1a}}{\longrightarrow} \mathcal{A}_1 + a \qquad \mathcal{B}_1 \overset{g_{1b}}{\longrightarrow} \mathcal{B}_1 + b$$

$$\mathcal{A}_0 \overset{g_{0a}}{\longrightarrow} \mathcal{A}_0 + a \qquad \mathcal{B}_0 \overset{g_{0b}}{\longrightarrow} \mathcal{B}_0 + b$$

$$a \overset{k_a}{\longrightarrow} \emptyset \qquad b \overset{k_b}{\longrightarrow} \emptyset$$

We describe the state of the system as the joint probabiity $P_{sa}{}_{sb}(n_a,n_b,t)$ of having at time t n_a proteins a, n_b proteins b, gene a in the s_a binding state (1 for unbound and 0 for bound) and gene b in the s_b binding state. We can then describe all possible transitions between all possible states of the system with a set of differential-difference equations describing the probability changes resulting from all possible transitions to and from any specific state at any point in time. Establishing an upper bound for the number of proteins N, it is possible to define a finite length vector $P(t)$ containing all of the probabilities of all possible states of the system at time t. This vector will be of size $4N^2$ for a network of 2 genes, a number that can be very large for systems having a large number of protein molecules. We can now describe the whole set of equations using a transition matrix A acting on the vector $\mathbf{P}(t)$. The transition matrix A has negative diagonal terms, representing the rates of exiting the states, and positive off-diagonal terms, representing the rates of transitions into those states. We have:

$$\frac{d\mathbf{P}(t)}{dt} = A\,\mathbf{P}(t) \qquad [1]$$

This equation can be used in a variety of ways, to provide numerically exact calculations of steady-state probability distributions, time-dependent probability distributions, or mean first-passage times. We will start by calculating steady-state probability distributions and a corresponding "effective potential" field. For this purpose we will use reflective boundary conditions in the vicinity of $n_a,n_b = N$, so that the transition matrix A acting on $\mathbf{P}(t)$ will conserve probability. Because all of the states in the toggle switch are interconnected, and because the matrix A conserves probability, there will be 1 and only 1 nondegenerate eigenvalue equal to zero. Since the steady-state probability distribution corresponds to the vector $\mathbf{P}_{ss}$ for which $d\mathbf{P}_{ss}/dt = 0$, finding the steady-state distribution means finding the first eigenvector of the transition matrix: $A\mathbf{P}_{ss} = 0\mathbf{P}_{ss}$. Using numerical methods to find the eigenvectors, we can obtain the steady-state probability distribution which can be used to calculate an "effective potential" field defined by $\mathbf{V}(n_a,n_b) \equiv -\log[\mathbf{P}(n_a,n_b)]$, plotted in Fig. 1 for the various switches.

To investigate how the system switches between the attractors, we can analyze a time-dependent solution obtained from Eq. **1**. This solution will depend on the exponential of the transition matrix and the initial state $\mathbf{P}_0$:

$$\mathbf{P}(t) = e^{At}\mathbf{P}_0 \qquad [2]$$

Choosing a desired time step Δt, we can calculate a matrix $\Delta A = e^{A\Delta t}$ and perform the calculations iteratively: $P(t + \Delta t) = P(t)\,\Delta A$. In Fig. 2, we plot probability distributions showing the escape mechanisms for different binding scenarios. For the binding of dimers, we see that the transitions occur through states with low levels of both proteins, indicating that the bound/bound binding state shuts down the production of both proteins to bring the system to a position where it can go to either attractor with equal probability. The exclusive switch does not allow the bound/bound binding state, so the transition occurs through the unbound/unbound binding state. Because the 2 protein species cannot be shut down at the same time, the switching trajectory does not go through states with simultaneously low levels of both proteins.

The λ-Phage. Aurell and Sneppen have found model parameters to describe the λ-phage switch. In the promoter region of the virus, 3 different operators can be bound by either of the 2 protein species involved, but with different binding affinities. Different DNA occupancy states will result in different rates of transcription of the 2 species. The binding affinities $G(s)$ of the transcription factors *cro* (a) and *cI* (b) to the DNA can be used to calculate the probabilities of promoter occupancy states s:

$$P_s = Z^{-1}[cI]^{i_s}[cro]^{j_s}e^{-G(s)/RT} \qquad [3]$$

$[cI]$ and $[cro]$ are concentration of dimers. i_s and j_s are the number of the respective dimers required for binding state s. Z is the partition function. The concentrations of dimers are calculated taking into account the dimerization equilibrium constants and cell volume. The average transcription initiation rates can now be calculated (0 = free, 1 = *cI*-bound, 2 = *cro*-bound):

$$g_a = R_{cro}(P_{000} + P_{100} + P_{200}) \qquad [4]$$

$$g_b = R_{cI1}(P_{010} + P_{011} + P_{012}) + R_{cI0}(P_{000} + P_{001} + P_{002} + P_{020} + P_{021} + P_{022}) \qquad [5]$$

R_{cro} is the transcription initiation rate of *cro* when RNA polymerase can bind the DNA, R_{cI1} is the activated transcription rate of *cI*, and R_{cI0} is its basal transcription rate. Considering x_a and x_b the number of protein molecules to result from a single transcription event (burst size), and k_a and k_b the degradation rates, the deterministic differential equations are:

$$\frac{da}{dt} = g_a(a,b)x_a - k_a a \qquad [6]$$

$$\frac{db}{dt} = g_b(a,b)x_b - k_b b \qquad [7]$$

The parameters used in these calculations are in Table S1. The essential binding affinities of the different DNA occupancy states are in Table S2.

Other DNA occupancy states resulting from adding 2 of the ones listed do not interact cooperatively, so their binding energy can be calculated simply by adding the binding energy of the two component states, e.g., $G(112) = G(110) + G(002)$.

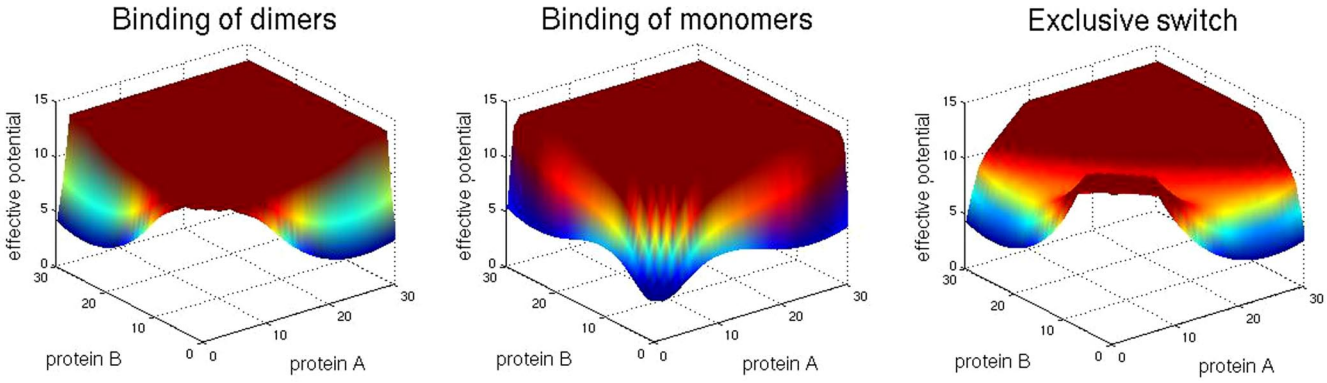

Fig. S1. Effective potential surfaces for different binding scenarios.

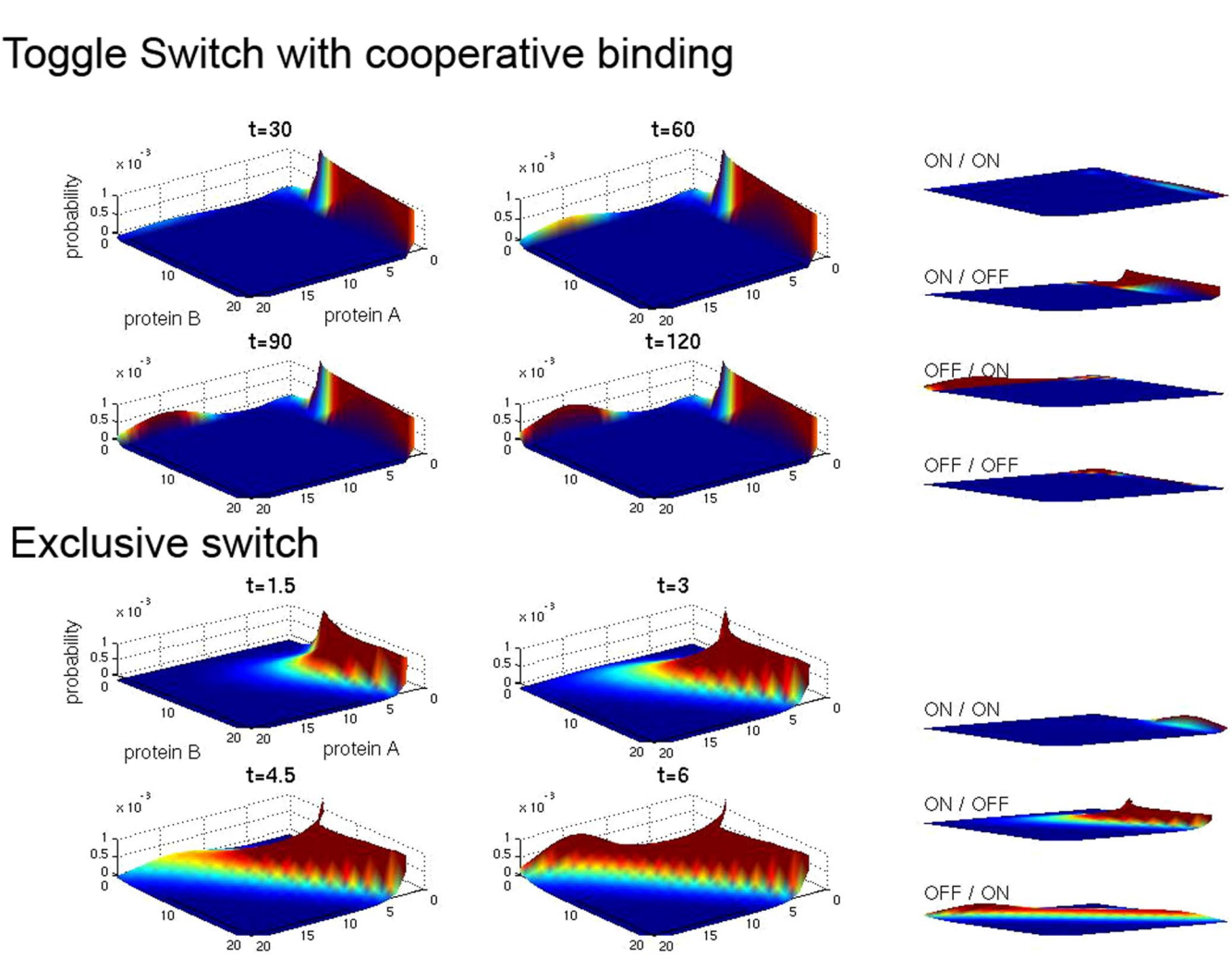

Fig. S2. Escape mechanisms for different binding scenarios. On the right, the final probability distribution is broken down into the different DNA occupancy states.

Table S1. Parameters used in Eqs. 6 and 7

Parameter	Value
Ideal gas constant	RT = 0.617 kcal/mol
Cell volume	V = $10^{-15}L$
cro burst size	$x_a = 20$
cI burst size	$x_b = 1$
cro transcr. init. rate	$R_{cro} = 0.30\ s^{-1}$
activated *cI* transcr. init. rate	$R_{cI1} = 0.115\ s^{-1}$
basal *cI* transcr. init. rate	$R_{cI0} = 0.01045\ s^{-1}$
cro degradation	$k_a = 0.0002\ s^{-1}$
cI degradation	$k_b = 0.00034\ s^{-1}$
cro dimerization constant	$K_{cro} = 2 \times 10^{-8}\ M$
cI dimerization constant	$K_{cI} = 3 \times 10^{-7}\ M$

transcr., transcription; init., initiation.

Table S2. Essential binding affinities of the different DNA occupancy states

State	Affinity, kcal/mol
000	0
001	−12.5
010	−10.5
100	−9.5
011	−25.7
110	−22.0
101	−22.0
111	−35.4
002	−14.4
020	−13.1
200	−15.5
220	−28.6
202	−29.9
022	−27.5
222	−43.0