

Stochastic Phenotype Transition of a Single Cell in an Intermediate Region of Gene State Switching

Hao Ge,^{1,2,*} Hong Qian,³ and X. Sunney Xie^{1,4,†}

¹*Biodynamic Optical Imaging Center (BIOPIC), Peking University, Beijing 100871, People's Republic of China*

²*Beijing International Center for Mathematical Research (BICMR), Peking University, Beijing 100871, People's Republic of China*

³*Department of Applied Mathematics, University of Washington, Seattle, Washington 98195, USA*

⁴*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 3 July 2014; revised manuscript received 20 November 2014; published 17 February 2015)

Multiple phenotypic states often arise in a single cell with different gene-expression states that undergo transcription regulation with positive feedback. Recent experiments show that, at least in *E. coli*, the gene state switching can be neither extremely slow nor exceedingly rapid as many previous theoretical treatments assumed. Rather, it is in the intermediate region which is difficult to handle mathematically. Under this condition, from a full chemical-master-equation description we derive a model in which the protein copy number, for a given gene state, follows a deterministic mean-field description while the protein-synthesis rates fluctuate due to stochastic gene state switching. The simplified kinetics yields a nonequilibrium landscape function, which, similar to the energy function for equilibrium fluctuation, provides the leading orders of fluctuations around each phenotypic state, as well as the transition rates between the two phenotypic states. This rate formula is analogous to Kramers' theory for chemical reactions. The resulting behaviors are significantly different from the two limiting cases studied previously.

DOI: 10.1103/PhysRevLett.114.078101

PACS numbers: 87.18.Cf, 82.40.Bj, 87.18.Tt

A single cell behaves stochastically with time as a consequence of gene expressions and biochemical regulations. The intrinsic stochasticity of cellular kinetics has two major origins: the stochastic gene state switching and copy-number fluctuations of proteins. The former is pertinent to the fact that there is only a single copy of DNA inside a typical cell that leads to stochastic productions of mRNA and protein [1], while the latter results from the low copy numbers of certain proteins [2]. According to the stochastic law of mass action, when a well-mixed ideal solution is at a chemical equilibrium, such as in a test tube, the copy-number distribution of proteins at the equilibrium steady state must have only a single peak [3], which is indicative of a sole phenotypic state. A living cell under a nonequilibrium steady state, which continuously exchanges materials and energies with its surroundings, however, usually can have multiple phenotypic states, corresponding to different modals of the copy-number distribution [4]. The coexistence of multiple phenotypic states, and transitions among them induced by intrinsic stochasticity, can be advantageous for the survival of cells in unpredictable environments [5]. Still, the maintenance of their stabilities and the transition rates among them are far from quantitatively understood; especially not enough attention has previously been paid to an intermediate region, in which the gene state switching is neither extremely slow nor extremely rapid. It is the case at least in *E. coli* according to recent experimental observations [1,4].

Here we consider the simplest gene network with a self-regulating protein, often referred to as a toggle switch [6], which consists of positive feedback as well as multiple gene states associated with different protein-synthesis rates [Fig. 1(a)]. We assume the protein functions as a dimer, which binds and activates its own gene. At a given moment of time, the chemical state of the cell is described by both the gene state $\{i = 1, 2\}$ and protein copy number $\{n = 0, 1, 2, \dots\}$ including those bound with the DNA molecule. In terms of chemical kinetics, the time evolution of the probability distribution of the chemical state (i, n) is governed by a chemical master equation (CME) [[7,8], Fig. 1(b)], i.e.,

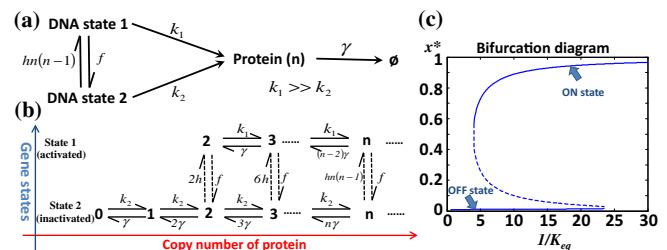


FIG. 1 (color online). (a) A minimal gene network with positive feedback and two different gene states. (b) The diagram of the full chemical master equation [see Eq. (1)]. (c) Deterministic mean-field model [Eq. (2)] with bistability induced by positive feedback, in which $k_1 = 10 \text{ min}^{-1}$, $k_2 = 0.1 \text{ min}^{-1}$, and $\gamma = 0.02 \text{ min}^{-1}$.

$$\begin{aligned}
\frac{\partial p_1(n, t)}{\partial t} &= k_1 p_1(n-1, t) - k_1 p_1(n, t) \\
&\quad + \gamma(n-1) p_1(n+1, t) - \gamma(n-2) p_1(n, t) \\
&\quad + hn(n-1) p_2(n, t) - f p_1(n, t), \\
\frac{\partial p_2(n, t)}{\partial t} &= k_2 p_2(n-1, t) - k_2 p_2(n, t) \\
&\quad + \gamma(n+1) p_2(n+1, t) - \gamma n p_2(n, t) \\
&\quad - hn(n-1) p_2(n, t) + f p_1(n, t), \quad (1)
\end{aligned}$$

in which k_1 and k_2 are the protein-synthesis rates for gene states 1 and 2, respectively, γ is the decay rate of the protein copy number that consists of the protein degradation as well as the cell division, and the switching rates between the two gene states are f and $hn(n-1)$.

We assume that the synthesis rate k_1 corresponding to the gene-active state (state 1) is sufficiently high while the k_2 associated with the gene-inactive state (state 2) is very low. Consequently, there emerge three time scales within this simple gene network: (i) the decay rate γ of the protein copy number, (ii) the switching rates f and $hn(n-1)$ between the gene states, and (iii) the larger protein-synthesis rate k_1 . Normally, the typical copy number of the protein when the cell is fully activated, k_1/γ , is quite high. This implies the time scale (iii) is usually much faster than (i). Most of the previous works have focused on two other scenarios: when (ii) is even much slower than (i), in which bimodal distribution of the protein copy number can occur even without positive feedback [9], or when (ii) is much faster than (iii), in which the gene states are in a rapid preequilibrium and often the diffusion approximation of CME is also applied [10,12]. Not enough attention has been paid to a third intermediate scenario in which the gene state switching is much faster than γ but actually is slower than the protein synthesis, thus the copy-number fluctuations of protein. This underexplored third scenario turns out to be most relevant for at least Lac operon, in which the stochastic kinetics of a single DNA molecule plays a rather significant role [1,4,13]. The ubiquitous transcriptional and translational bursts observed in living cells recently ranging from bacterial to mammalian cells also indicate the relatively slow switching between the active and inactive states of genes [1,14]. The transition rate in the last case was studied in [15] by an intuitive approach.

In the present Letter, we derive a much simpler stochastic model from the full CME of the gene-regulation network in Fig. 1(a) for this third intermediate region, which is easier to handle. We further propose a saddle-crossing rate formula between the two phenotypic states, together with an emerging landscape function that is the analog of the energy function in the nonequilibrium case. Further, we also show that the resulting behaviors can be very different from other limiting cases studied previously.

When both the fluctuations within gene state switching and evolution of the protein copy number are extremely

rapid, the gene states are at a rapid preequilibrium and a rescaled protein dynamics follows the mean-field rate equations in terms of a continuous variable x , the ratio of the protein copy number n to $n_{\max} = k_1/\gamma$ [[10,16], Fig. 1(c)], i.e.,

$$\frac{dx}{dt} = g(x) - \gamma x, \quad (2)$$

in which the dynamically averaged protein-synthesis rate $g(x) = [\bar{h}x^2 k_1 + f k_2/n_{\max}(\bar{h}x^2 + f)] = [\gamma(x^2 + K_{eq} k_2/k_1)/x^2 + K_{eq}]$, $K_{eq} = f/\bar{h}$, and $\bar{h} = hn_{\max}^2$. The mean-field dynamics is really macroscopic dynamics, not mean dynamics of a mesoscopic system.

In the presence of positive feedback, the dynamics, Eq. (2), can have two stable fixed points [the off and on states in Fig. 1(c)] separated by an unstable one at certain range of biochemical environmental parameters. A phase diagram is usually employed to precisely characterize the complete range of environmental parameters over which the system is bistable [17], Fig. 1(c)]. The system will undergo a switching from bistability to monostability and vice versa at certain critical environmental parameter values [Fig. 1(c)].

We further assume that the gene state switching is much slower than the active protein synthesis but much faster than cell division in this intermediate region that most relevant to the real situation in at least *E. coli*. Under this condition, the copy-number fluctuation of the protein can be safely considered to be nearly neglectable, and the full CME in Fig. 1(a) can be reduced to a much simpler model, called single-molecule fluctuating-rate model, in which the protein copy number given each gene state follows the deterministic mean-field description but the protein-synthesis rates are fluctuating due to stochastic gene state switching [Fig. 2(a) [10]]. A similar fluctuating-rate model has appeared in single-molecule enzyme kinetics [18]. Note that such a simplified model is also valid for the case in which the gene state switching is extremely slow, even much lower than the cell division.

The stochastic dynamics of the fluctuating-rate model can be simulated by the Doob-Gillespie method [8,10,19], and the time evolution of the probability $p_i(x)$ of the cell state (i, x) is described by [[10], Fig. 2(a)]

$$\begin{aligned}
\frac{\partial p_1}{\partial t} &= -f p_1 + \bar{h} x^2 p_2 - \frac{\partial}{\partial x} \left[\left(\frac{k_1}{n_{\max}} - \gamma x \right) p_1 \right], \\
\frac{\partial p_2}{\partial t} &= f p_1 - \bar{h} x^2 p_2 - \frac{\partial}{\partial x} \left[\left(\frac{k_2}{n_{\max}} - \gamma x \right) p_2 \right]. \quad (3)
\end{aligned}$$

On the other hand, as in many of the previous works [12], it is often assumed that the gene state switching is extremely rapid. In this case, the full CME in Fig. 1(b) can be reduced to a different simplified model [Fig. 2(c)] [10]: simply integrating all the gene states that are at fast equilibrium. The time evolution of the protein copy-number distribution $p(n, t)$ is [10]

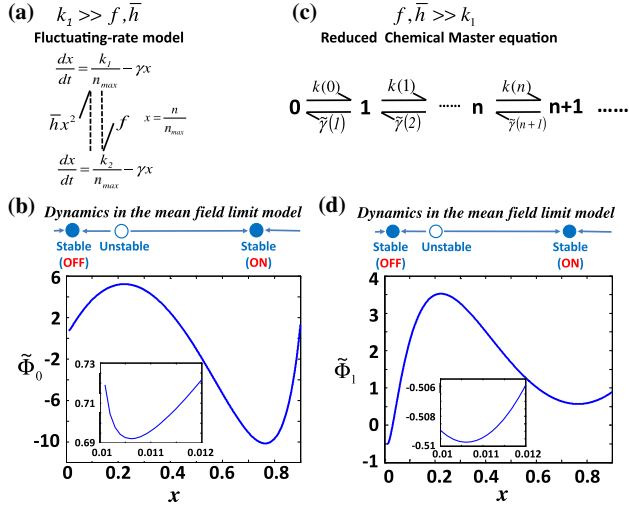


FIG. 2 (color online). In the intermediate region of gene state switching, the full CME can be simplified to a fluctuating-rate model (a), the steady-state distribution of which corresponds to a normalized landscape function $\tilde{\Phi}_0(x)$ (b). In the case if the gene state switching is extremely rapid, a reduced CME (c) and a different landscape function $\tilde{\Phi}_1(x)$ (d) can also be derived. The insets in (b) and (d) are the zoom in of the functions near $x = 0$. The parameters in (b) and (d) are the same as those in Fig. 1(c) with $K_{eq} = 1/5.5$.

$$\frac{\partial p(n, t)}{\partial t} = k(n-1)p(n-1, t) - k(n)p(n, t) + \tilde{\gamma}(n+1)p(n+1, t) - \tilde{\gamma}(n)p(n, t), \quad (4)$$

in which $k(n) = [(k_1 hn(n-1) + k_2 f/hn(n-1) + f)]$ is the fast-equilibrated protein-synthesis rate, and $\tilde{\gamma}(n) = [hn(n-1)(n-2) + fn/hn(n-1) + f]\gamma$ is the fast-equilibrated protein decay rate.

For each of the two simplified models, we can derive a nonequilibrium landscape function of x from the WKB method [10,16,20], approximating the negative logarithm of the stationary distribution $p^{ss}(x)$ of x as the noise is relatively small. The landscape function $\Phi_0(x)$ associated with the fluctuating-rate model [Fig. 2(a)] satisfies

$$\frac{d\Phi_0(x)}{dx} = \frac{f}{\frac{k_1}{n_{\max}} - \gamma x} + \frac{\bar{h}x^2}{\frac{k_2}{n_{\max}} - \gamma x}, \quad (5)$$

and the landscape function $\Phi_1(x)$ for the reduced CME model [Fig. 2(c)] satisfies

$$\frac{d\Phi_1(x)}{dx} = -n_{\max} \log \frac{g(x)}{\gamma x}. \quad (6)$$

See [10] for the detailed derivation. Nonequilibrium here means these landscape functions are not the potential of the right-hand side of the mean-field model, Eq. (2), and also the detailed balance is broken in the CME description [Fig. 1(b)].

These landscape functions for a living cell are generalizations of the energy landscapes widely employed in nondriven biochemical systems such as protein folding [21]. Distinctly contrary to the latter, a nonequilibrium landscape function is not given *a priori* to a dynamical system; it is actually an emergent consequence from the detailed chemical kinetics.

Notice that the mean-field dynamics, Eq. (2), depends on three independent parameters given the unit of time, i.e., γ , K_{eq} , and k_1/k_2 . Once they are given, each landscape function still depends linearly on one more parameter: a scalar multiplier. Hence, we can define the normalized landscape functions $\tilde{\Phi}_0(x) = \Phi_0(x)/f$ and $\tilde{\Phi}_1(x) = \Phi_1(x)/k_1$ [Figs. 2(b) and 2(d)], which also only depend on the three independent parameters.

The most important feature of these landscape functions is that the corresponding deterministic mean-field dynamics in Eq. (2) always goes downhill [10,20,22,23]. This implies that any local minimum (maximum, saddle) of a landscape function corresponds to a stable (unstable, saddle) steady state of the respective deterministic mean-field dynamics, Eq. (2) [Figs. 2(b) and 2(d)]. It also implies the parameter ranges for double-well shaped landscape functions are the same for both $\tilde{\Phi}_0(x)$ and $\tilde{\Phi}_1(x)$. Furthermore, the variance of local fluctuations around each stable steady state x^* (phenotypic state) can be approximated by $1/[d^2\Phi_i(x)/dx^2]_{x=x^*}$ [10,20]. One can clearly see from Figs. 2(b) and 2(d) that the local fluctuation in the intermediate region can be very different from that in the case with extremely rapid gene state switching, even if the mean-field model is kept the same.

It is indispensable to emphasize that although the diffusion approximation of CME that was always applied [12,15] can also give rise to a landscape function, in the worst-case scenario, it might even reverse the relative stability of the coexisting phenotypic states and give incorrect saddle-crossing rates [10,24].

In general, the transition rate between phenotypic states is defined as the reciprocal of the mean-first-passage time starting from one phenotypic state to another [25]. It is well defined because the mean-first-passage time is nearly independent of the initial values within a same phenotypic state, as long as there is a time-scale separation of intraphenotype fluctuations and interphenotype transitions.

The stochastic gene state switching can be the rate-limiting step for the transition between phenotypic states in case it is extremely slow [15]. However, the rate formulas can be much more complicated in the intermediate region as well as the extremely rapid region.

In the latter two cases, following the mathematical derivation in the Freidlin-Wentzell large deviation theory and related other theoretical works in physics and chemistry [22,26], the transition rates from one phenotypic state A transiting to the other phenotypic state B can be expressed as

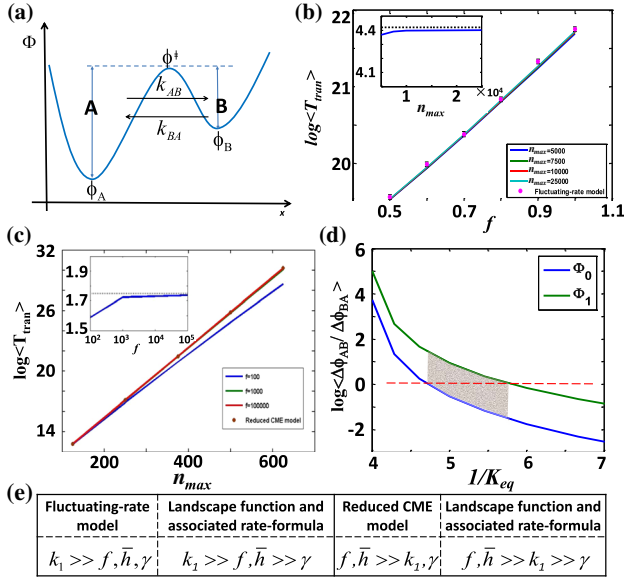


FIG. 3 (color online). The general saddle-crossing rate formulas $k_{AB} \approx k_{AB}^0 \exp(-\Delta\Phi_{AB})$ (a) and the mean transition time from the off state to the on state in the full CME and both simplified models (with different colors) are obtained as either the gene state switching is in the intermediate region (b) or the extremely rapid region (c). All the simulated data are obtained through the mean-first-passage-time method, except that from the fluctuating-rate model which is obtained through stochastic simulation of the trajectories. (d) The relative stability of the two phenotypic states with respect to Φ_0 (blue) and Φ_1 (green) as a function of $1/K_{eq}$. The multicolored lines in (b) and (c) almost overlap with each other. The insets in (b) and (c) compare the numerically determined normalized barrier heights from the full CME (solid blue) and the two simplified models (dashed black). The parameter $(k_1/k_2) = 3000$ in (b), 50 in (c), and 100 in (d). $K_{eq} = 1/6$ in both (b) and (c). $\gamma = 0.02 \text{ min}^{-1}$. (e) Summarizing the conditions under which the simplified models and associated landscape functions as well as saddle-crossing rate formulas are valid.

$$k_{AB} \approx k_{AB}^0 \exp(-\Delta\Phi_{AB}), \quad (7)$$

where $\Delta\Phi_{AB}$ is called the barrier term and k_{AB}^0 is a prefactor with unit time^{-1} only dependent on the three parameters in the mean-field model. The barrier $\Delta\Phi_{AB}$ of the landscape function $\Phi(x)$ is $\Delta\Phi_{AB} = \Phi^\ddagger - \Phi_A$, where Φ^\ddagger is the landscape value of the unstable fixed point (local maximum) along the transition path from the phenotypic state A to another state B, and Φ_A is the landscape value of the phenotypic state A (local minimum) [Fig. 3(a)]. Note that the expressions of the two landscape functions, Eqs. (5) and (6), as well as the saddle-crossing rates, Eq. (7), in such a simple case can be straightforwardly solved here.

This formula is quite similar to the well-known Kramers' formula as well as the Arrhenius equation for the temperature dependence of the reaction rate [27], and the barrier term $\Delta\Phi_{AB}$ is an analog of activation energy. The exponential dependence of the landscape barrier in such a

saddle-crossing rate formula, Eq. (7), directly guarantees the strong stability of phenotypes against intrinsic stochasticity. Our analytical theory is consistent with previous works on the phenotypic-state transition, which were based on numerical simulations [28], or ideas from the transition-state theory [15,29].

For the fluctuating-rate model [Fig. 2(a)], the protein synthesis occurs in bursts. Once the corresponding steady-state value x_{off} of the off state is extremely low, the barrier height in our rate formula can be approximated by $((x_{\text{trans}} - x_{\text{off}})/b)$, where x_{trans} is the value of x at the barrier and b is the burst size [10]. Such an approximated barrier height is the same as the rate formula proposed in [30] for bursty dynamics as well as that in the nonadiabatic rate theory [15]. Recently, Assaf *et al.* and Lv *et al.* [16] have also proposed a saddle-crossing rate formula for the case when the two time scales, the gene state switching and protein copy-number fluctuation, are comparable.

We performed numerical simulations to validate the rate formula, Eq. (7), by either simulating stochastic trajectories or numerically solving the equations of mean-first-passage time [10,25,31], obtaining the mean transit time $\langle T_{\text{tran}}^{\text{off} \rightarrow \text{on}} \rangle$ from the off state to the on state. We keep the mean-field dynamics unchanged, i.e., fixing the parameters K_{eq} , k_1/k_2 , and γ , and letting f and n_{max} vary. The results illustrate that $\langle T_{\text{tran}}^{\text{off} \rightarrow \text{on}} \rangle$ in the full CME model is well approximated by the two simplified models in their separate regions of gene state switching [Figs. 3(b) and 3(c)], and the normalized barriers $\Delta\tilde{\Phi}_{\text{off},\text{on}}$ with respect to f or n_{max} can be determined from such Arrhenius-like plots [Figs. 3(b) and 3(c) inset], which matches those predicted from the normalized landscape functions, $\tilde{\Phi}_0(x)$ or $\tilde{\Phi}_1(x)$. Figures 3(b) and 3(c) can be experimentally observed once we can simultaneously tune at least two of the parameters, and keep the mean-field model unchanged [5]. In such types of experiments, we will see that in the intermediate region, $\langle T_{\text{tran}}^{\text{off} \rightarrow \text{on}} \rangle$ can be insensitive to the total number of protein molecules [Fig. 3(b)], but varies dramatically on the gene state switching rates; while in the extremely rapid region, the situation is just opposite [Fig. 3(c)]. The conditions under which the simplified models and associated landscape function, as well as saddle-crossing rate formulas, being valid, are summarized in Fig. 3(e).

The complex parameter-dependence relation implies that the relative stability between the two phenotypic states, defined as the ratio of the two ‘‘activation energies’’ $\Delta\Phi_{AB}$ and $\Delta\Phi_{BA}$ might be reversed, as the gene state switching rates is in the intermediate region or the extremely rapid region, even keeping the same equilibrium constants among all the gene states [see the shaded region in Fig. 3(d), and also comparing Figs. 2(b) and 2(d)].

The results in the present Letter can be similarly generalized to any self-activating regulatory module. See [10].

As a conclusion, recent single-molecule experiments revealed that the stochastic gene state switching in a living

cell is possibly not sufficiently rapid to meet the requirement of the rapid-equilibrium assumption, but also is not slow enough to meet the nonswitching assumption. In the present Letter, we proposed a simplified model for the intermediate scenario, which is significantly simpler than the full CME description. A nonequilibrium landscape function and an associated saddle-crossing rate formula, in the similar form as Kramers' formula, for the phenotype switching are derived. Even with the simplest module of gene-regulation networks, we show that the rapid-equilibrium assumption can result in very different behaviors. Our theory indicates that the stochastic nature of a single DNA molecule is essential for discrete phenotypic cellular states and their functions.

We thank P. Choi and T.-J. Li for discussion. H. Ge is supported by NSFC 10901040, 21373021, and the Foundation for Excellent Ph.D. Dissertation from the Ministry of Education in China (No. 201119). X. S. Xie is supported by the NIH Pioneer Grant (No. 1DP1OD000277).

*haoge@pku.edu.cn

†xie@chemistry.harvard.edu

- [1] G. W. Li and X. S. Xie, *Nature (London)* **475**, 308 (2011); Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, *Science* **329**, 533 (2010).
- [2] A. Eldar and M. Elowitz, *Nature (London)* **467**, 167 (2010).
- [3] D. Shear, *J. Chem. Phys.* **48**, 4144 (1968).
- [4] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, *Science* **322**, 442 (2008); P. B. Gupta, C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser, and E. S. Lander, *Cell* **146**, 633 (2011); E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden, *Nature (London)* **427**, 737 (2004).
- [5] E. Kussell and S. Leibler, *Science* **309**, 2075 (2005); M. Acar, J. T. Mettetal, and A. van Oudenaarden, *Nat. Genet.* **40**, 471 (2008).
- [6] T. S. Gardner, C. R. Cantor, and J. J. Collins, *Nature (London)* **403**, 339 (2000).
- [7] M. Delbrück, *J. Chem. Phys.* **8**, 120 (1940); R. Grima, D. R. Schmidt, and T. J. Newman, *J. Chem. Phys.* **137**, 035104 (2012).
- [8] D. T. Gillespie, *J. Comput. Phys.* **22**, 403 (1976); *J. Phys. Chem.* **81**, 2340 (1977).
- [9] R. Karmakar and I. Bose, *Phys. Biol.* **1**, 197 (2004); M. N. Artyomov, J. Das, M. Kardar, and A. K. Chakraborty, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18958 (2007); H. Qian, P. Z. Shi, and J. Xing, *Phys. Chem. Chem. Phys.* **11**, 4861 (2009); T. To and N. Maheshri, *Science* **327**, 1142 (2010); J. E. M. Hornos, D. Schultz, G. Innocentini, J. Wang, A. Walczak, J. Onuchic, and P. Wolynes, *Phys. Rev. E* **72**, 051907 (2005); H. Feng, B. Han, and J. Wang, *J. Phys. Chem. B* **115**, 1254 (2011); A. Ochab-Marcinek and M. Tabaka, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22096 (2010).
- [10] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.114.078101>, which includes Ref. [11], for the detailed derivation of all the equations and the numerical method for simulation.
- [11] T. G. Kurtz, *J. Chem. Phys.* **57**, 2976 (1972).
- [12] X.-M. Zhu, L. Yin, L. Hood, and P. Ao, *Funct. Integr. Genomics* **4**, 188 (2004); J. Wang, L. Xu, E. Wang, and S. Huang, *Biophys. J.* **99**, 29 (2010); J. X. Zhou, M. D. S. Aliyu, E. Aurell, and S. Huang, *J. R. Soc. Interface* **9**, 3539 (2012); H. Qian, *Quant. Biol.* **2**, 47 (2014); M. Y. Lu, J. Onuchic, and E. Ben-Jacob, *Phys. Rev. Lett.* **113**, 078102 (2014).
- [13] P. Bhogale, R. Sorg, J.-W. Veening, and J. Berg, *Nucleic Acids Res.* **42**, 11321 (2014).
- [14] D. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, *Science* **332**, 472 (2011).
- [15] A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18926 (2005).
- [16] M. Assaf, E. Roberts, and Z. Luthey-Schulten, *Phys. Rev. Lett.* **106**, 248102 (2011); C. Lv, X. Li, F. Li, and T. Li, *PLoS One* **9**, e88167 (2014).
- [17] J. Ferrell and W. Xiong, *Chaos* **11**, 227 (2001).
- [18] W. Min, I. V. Gopich, B. P. English, S. C. Kou, X. S. Xie, and A. Szabo, *J. Phys. Chem. B* **110**, 20093 (2006).
- [19] J. L. Doob, *Trans. Am. Math. Soc.* **58**, 455 (1945).
- [20] H. Ge and H. Qian, *Phys. Rev. Lett.* **103**, 148103 (2009).
- [21] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, *Science* **254**, 1598 (1991); J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
- [22] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Spinger, New York, 1998).
- [23] H. Ge and H. Qian, *Chaos* **22**, 023140 (2012).
- [24] M. Vellela and H. Qian, *J. R. Soc. Interface* **6**, 925 (2009).
- [25] K. Schulten, Z. Schulten, and A. Szabo, *J. Chem. Phys.* **74**, 4426 (1981).
- [26] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, *J. Chem. Phys.* **100**, 5735 (1994), and references therein.
- [27] P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- [28] P. B. Warren and P. R. ten Wolde, *Phys. Rev. Lett.* **92**, 128101 (2004); *J. Phys. Chem. B* **109**, 6812 (2005); P. Mehta, R. Mukhopadhyay, and N. S. Wingreen, *Phys. Biol.* **5**, 026005 (2008).
- [29] T. Ushikuboa, W. Inoueb, M. Yodaa, and M. Sasai, *Chem. Phys. Lett.* **430**, 139 (2006).
- [30] P. J. Choi, X. S. Xie, and E. I. Shakhnovich, *J. Mol. Biol.* **396**, 230 (2010).
- [31] P.-Z. Shi and H. Qian, *J. Chem. Phys.* **134**, 065104 (2011).