Two-Component Genetic Switch as a Synthetic Module with Tunable Stability

C.-M. Ghim[1] and E. Almaas[1,2,*]

[1]*Biosciences & Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[2]*Department of Biotechnology, Norwegian University of Science and Technology, Trondheim, Norway*

(Received 17 November 2008; published 6 July 2009)

Despite stochastic fluctuations, some genetic switches are able to retain their expression states through multiple cell divisions, providing epigenetic memory. We propose a novel rationale for tuning the functional stability of a simple synthetic gene switch through protein dimerization. Introducing an approximation scheme to access long-time stochastic dynamics of multiple-component gene circuits, we find that the spontaneous switching rate may exhibit greater than 8 orders of magnitude variation. The manipulation of the circuit's biochemical properties offers a practical strategy for designing robust epigenetic memory with synthetic circuits.

DOI: 10.1103/PhysRevLett.103.028101 PACS numbers: 87.18.Cf, 05.40.Ca, 82.39.−k

One of the most significant challenges in cell biology of the postgenome era is to understand how the functional repertoire of cells is related to system-level properties of complex signaling networks. The suggestion that cellular networks are organized around functional modules [1] has led the way for large-scale studies of their organizing principles [2]. A particularly important class of functional modules is able to behave as bistable switches, providing a cell with the ability to change between two discrete behavioral states. In general, the bistable behavior allows cells to retain information about transient intracellular signals for many generations [3], thus serving as epigenetic "memory" devices. For instance, the lambda phage toggle is capable of sustaining its lysogenic reproduction mode for $\sim 10^7$ generations before spontaneously exiting [4]. This is an impressive feat, as transitions between the different states of a toggle can be driven by fluctuations in the abundance of its constituting proteins. Furthermore, by tuning the toggle's spontaneous switching rate, it is possible to generate phenotypic diversity in populations that matches fluctuations in nutrient availability [5], thus providing a directly measurable fitness effect.

Recently a simple toggle switch consisting of a pair of transcriptionally repressing genes was constructed [6], demonstrating the feasibility of *de novo* synthesis of cellular memory units. Scalable strategies for integrating the toggle as part of gene circuits with more elaborate functionalities, however, critically relies on the quantitative understanding of the toggle's capacity to generate robust yet tunable switching behavior. Here we identify key factors that modulate the stability of various genetic toggle switches and propose a novel method that is based on the manipulation of fast binding-unbinding dynamics among proteins, DNA, and other macromolecules.

To quantitatively evaluate the toggle's robustness against random fluctuations and its dependence on circuit topology and kinetic details, we use the chemical master equation. We have identified experimentally tunable parameters and evaluated their effects on the switching rate. For biologically relevant parameter choices [7], we find that the switching rate can be tuned over a range of more than 8 orders of magnitude, reaching ultrastability. We discuss its evolutionary implications and hypothesize that our approach is generally applicable for introducing tunable stability in engineered biological circuits.

Deterministic analysis of model circuits.—The network of two mutually repressing genes "1" and "2" is schematically shown in Fig. 1. In this Letter, we have explored three reaction networks corresponding to distinct post-translational binding dynamics [Fig. 1(b)]: (i) HET circuit allows the proteins to form functionally inactive heterodimers; (ii),(iii) HOM1 (HOM2) circuit allows each protein species to form homodimers and the monomer (dimer) is the active form of the repressor. The full set of reactions are listed in Table I (see also supplemental files [8]).

Since the reversible binding reactions [Fig. 1(c)] equilibrate very fast in the physiological time scales for synthesis and degradation of macromolecules [9], we may character-

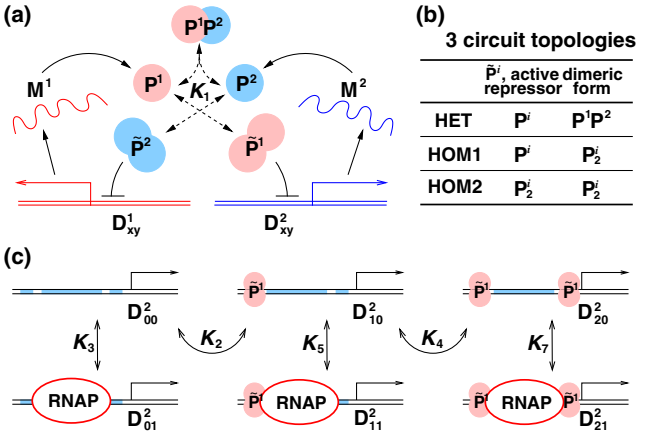

FIG. 1 (color online). (a) Circuitry of genetic toggle switch. (b) Three circuits studied. (c) Schematic of DNA binding states for gene 2 (similar for gene 1). D^i_{xy} is the binding status of the operator and promoter sequence of gene i. RNAP, RNA polymerase; $\tilde{P}^i$, active repressor; K_i, dissociation constant.

TABLE I. The full list of reactions for gene $i = 1, 2$ that are shared by different toggle switches (upper panel) and are model specific (lower panel). E^i denotes the DNA-RNAP complex with RNAP cleared of the *cis*-regulatory region.

	Reaction	Rate	Reaction	Rate
	$D^i_{00} + \tilde{P}^{\varepsilon_i} \rightleftharpoons D^i_{10}$	(k_2, q_2)	$D^i_{*1} \to E^i + D^i_{*0}$	α_m
	$D^i_{10} + \tilde{P}^{\varepsilon_i} \rightleftharpoons D^i_{20}$	(k_4, q_4)	$E^i \to M^i + E^i$	α'_m
Common	$D^i_{00} + \text{RNAP} \rightleftharpoons D^i_{01}$	(k_3, q_3)	$M^i \to P^i + M^i$	α_p
	$D^i_{10} + \text{RNAP} \rightleftharpoons D^i_{11}$	(k_5, q_5)	$M^i \to \varnothing$	γ_m
	$D^i_{20} + \text{RNAP} \rightleftharpoons D^i_{21}$	(k_7, q_7)	$P^i \to \varnothing$	γ_p
HET	$P^1 + P^2 \rightleftharpoons P^1P^2$	(k_1, q_1)	$P^1P^2 \to \varnothing$	γ_p/σ
HOMx	$P^i + P^i \rightleftharpoons P^i_2$	(k_1, q_1)	$P^i_2 \to \varnothing$	γ_p/σ

ize each binding-unbinding reaction pair by a single thermodynamic constant $K_i = k_i/q_i$, where k_i and q_i are the association and dissociation rates (Table I). Neglecting the effects of stochastic fluctuations, the reaction network in a steady state can be reduced to

$$\dot{p}_i = F_i(p_i, p_{\varepsilon_i}) - G_i(p_i, p_{\varepsilon_i}), \qquad i = 1, 2, \tag{1}$$

where the interaction between the two genes is reflected in the regulated synthesis rate F_i via $\varepsilon_i = \text{mod}(i, 2) + 1$, and G_i accounts for loss due to reversible binding kinetics as well as dilution due to cell growth. Note that p_i is the dimensionless concentration of the *active* repressor protein, being either P^i itself (HET and HOM1) or the homodimer P^i_2 (HOM2). For analytical simplicity, we explore the symmetric circuit.

Instead of representing the regulated synthesis rate F_i as an *ad hoc* sigmoid with an arbitrary Hill coefficient, we derived its functional form directly from the statonarity condition for the whole set of binding-unbinding reactions. Without loss of generality, we allow for the possibility of RNAP binding to the promoter site despite steric hindrance by operator-bound repressors; i.e., the dissociation constants K_5 and/or K_7 in Fig. 1(c) are finite. F can be factorized as $F(x, y) = \varphi(x)f(y)$, where $\varphi(x)$ is either 1 (HET, HOM1) or $2\sqrt{\theta x}$ (HOM2) with $\theta = K_2/K_1$, and the resulting form of the regulated synthesis rate is

$$f(y) = \lambda\left(1 + \frac{\nu}{1 + \mu y + \mu r y^2}\right), \tag{2}$$

where $\lambda = f(\infty)$ is the constitutive synthesis rate of proteins in full repression, $\nu = [f(0) - f(\infty)]/f(\infty)$ is the reduced fold change between on and off states. We introduce the cooperativity in repressor-DNA binding by allowing the active repressor to bind two binding sites with occupation-dependent binding affinity ($r > 1$). Notice, however, that cooperativity is not a necessary condition for the toggle behavior [10]. Following mass-action kinetics, the dimensionless parameters can be represented by equilibrium constants:

$$\lambda = \frac{\beta}{1 + u/s}, \qquad \mu = \frac{s + u}{1 + u}, \qquad \nu = \frac{u(1 - s)}{s(1 + u)}, \qquad r = \frac{K_2}{K_4},$$

where $s = K_3/K_5 \in (0, 1)$ is the promoter leakage which plays the role of inverse coupling strength between the two genes, and $u = K_3/[\text{RNAP}]$ is the RNAP-promoter dissociation constant scaled by the concentration of free RNAP. The gene expression efficiency is $\beta = \alpha_m\alpha_p/K_2\gamma_m\gamma_p$, where α_m (α_p) and γ_m (γ_p), respectively, denote the synthesis and degradation rate mRNA (protein). The decay term $G(x, y)$ is given by $x + xy/\kappa$ (HET), $x + 2x^2/\kappa$ (HOM1), or $2(x + 2x^{3/2}/\kappa\theta)$ (HOM2), where $\kappa = \sigma/\theta$ and σ is the ratio of dimer to monomer lifetime. We have used the experimentally determined reaction rates for the phage λ switch [7], and examined the dynamical behavior of each circuit in a wide range around these parameter values. Unless otherwise noted, $s = 0.01$, $\beta = 17.5$, $u = 3$, and $r = 25$.

Figure 2 shows the parameter regions for multistability identified by null cline analysis. We find that the dissociation constant of the protein dimer K_1, the promoter leakage s, and the gene expression efficiency β are key parameters with direct biological interpretation that control circuit stability. Regardless of circuit topology, a decrease in the promoter leakage s moves the system dynamics deeper inside the bistable region. In the limit of $K_1 \to \infty$, both HET and HOM1 circuits converge to the simplest monomer-only (MO) model, where the monomer is the exclusive form of the repressor protein and no interprotein binding is allowed. The convergence is already sufficient at $K_1 = 1000$ nM, for which the phase boundaries of HET and HOM1 coincide. A decrease in K_1 monotonically enhances the bistable region in the HET circuit while shrinking it for HOM1. For the HOM2 circuit, an increase in K_1 approximately parallel shifts the bistability region to higher values of β.

Interestingly, in the strong-binding regime combined with high expression efficiency and intermediate promoter leakage (wedge-shaped region in Fig. 2), HET develops higher-order multistability that harbors a *coexistence* state ($p_1 \sim p_2$) in addition to the two *monopoly* states ($p_1 \gg p_2$ or $p_2 \gg p_1$). Unlike the HOMx circuits, protein dimerization in HET acts as an extra coupling between the genes: The majority protein not only represses the transcription of the minority species, but also reduces its own repression by the minority species through binding the minority protein.

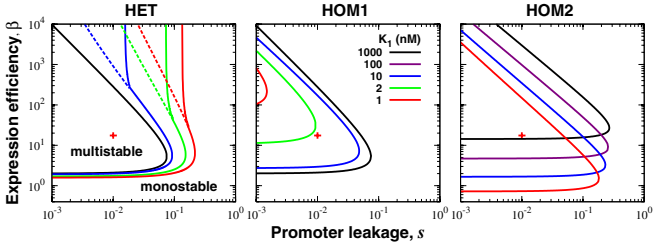

FIG. 2 (color online). Phase diagram of toggle circuits. Red + ($s = 0.01$, $\beta = 17.5$) indicates our base parameter choice. The HET circuit develops 3 or more stable states for large values of β and intermediate s (wedge-shaped area bounded by solid and broken lines).

This "rich-gets-richer" mechanism enhances the stability of steady states in the otherwise MO system while allowing the acquisition of various higher-order multistability states. Transitions between different stability regions belong to the supercritical (subcritical) pitchfork or saddle-node bifurcation (see the null cline analysis in [8]).

Stochastic dynamics.—We used the improved Gillespie method (Gibson-Bruck algorithm [11]) to generate exact time series for the full reaction system. We first qualitatively address the tunability of each circuit by studying the effective *potential* landscape $V_{\text{eff}} = -\ln P(n_1 - n_2)$, where n_i is the copy number of the active repressor and $P(x)$ is the relative frequency or probability of observing $n_1 - n_2 = x$. The double-well feature of Fig. 3 supports our deterministic bifurcation analysis. For the HET circuit, an increase in the dimer affinity (reduced K_1) moves the system deeper inside the bistability region (Fig. 2) and leads to a monotonic increase in barrier height. In contrast, the stability of HOM1 and HOM2 is mainly determined by the separation of the attractors. Figure 3 also demonstrates that an increase in the average number of proteins synthesized per single mRNA molecule or burst size $b = \alpha_p/\gamma_m$ (while keeping β unchanged) quite effectively lowers the potential barrier between the two dynamical states in all of the circuits. In passing, we note the appearance of a stochastically induced steady state, $n_1 \approx n_2$ for HOM1 with $b = 1$ and for HET with $b = 4$ [12].

The computational cost of exact simulations of the full reaction set is prohibitive, necessitating an approximate stochastic representation of the toggle circuit for studies of long-time processes such as switching (see, e.g., Allen *et al.* [13] for an alternative approach). We develop an approximate stochastic model of the circuit activity using a birth-death process of interacting proteins where we use the propensities for synthesis and decay given in Eq. (1). The resulting model thus corresponds to the average rates

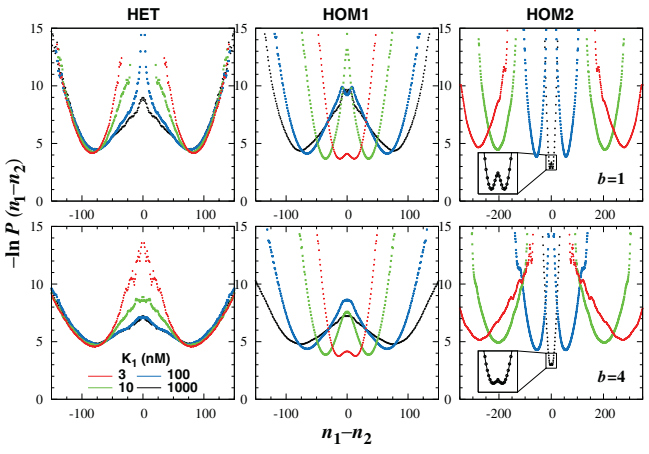

FIG. 3 (color). The effective potential for a bistable switch with $\beta = 17.5$ and $s = 0.01$. Each curve was generated by simulations of the full reaction set before being symmetrized. The upper (lower) row corresponds to the burst size of $b = 1$ ($b = 4$).

derived under the adiabatic approximation. To explicitly address the important effects of translational bursting [14], we factorize the average birth rate into the product of the stoichiometric coefficient (b) and the propensity. We assume that both mRNA and protein synthesis follow a Poisson process, which may not hold in the case of transcriptional bursting which is particular for eukaryotic cells [15]. The resulting master equation for the probability of the proteins' copy number configuration $\boldsymbol{n} = (n_1, n_2)$ reads

$$\dot{\mathcal{P}}(\boldsymbol{n},t) = \sum_{i=1,2} (\hat{L}_{\varepsilon_i}^{b} - 1)F(\boldsymbol{n})\mathcal{P}(\boldsymbol{n},t) + \sum_{i=1,2} (\hat{R}_i - 1)G(\boldsymbol{n})\mathcal{P}(\boldsymbol{n},t), \tag{3}$$

where the raising (lowering) operator $\hat{R}_i$ ($\hat{L}_i$) acts on species i. This reduced master equation can be translated into a Markov process governed by the imaginary-time Schrödinger equation which we solve through matrix diagonalization. Note that the lowest eigenvalue E_0 is zero due to its Laplacian nature and corresponds to the stationary configuration. While higher eigenmodes govern fast relaxations, the smallest nonzero eigenvalue ($E_1 > 0$) characterizes the rate of switching between the two stable states. Since the spectral gap $\Delta = E_2 - E_1$ is much bigger than E_1, we can represent an arbitrary state as a linear combination of the ground state and the first-excited state [16], effectively defining a two-state model where the interstate transition rate is solely determined by E_1.

Figure 4(a) shows a heat map of the switching rates for the three toggle circuits calculated using Eq. (3). A comparison with switching rates obtained from exact simulation of the full reaction set demonstrates that our approximation is highly accurate [Figs. 4(b) and 4(c)]. While we observe that enhanced stability with reduced s and large β is a common trend for the three circuits, they display significant differences in their tunability. In particular, a comparison with the MO system [dashed line in Figs. 4(b) and 4(c)] supports our inference from Fig. 3: HET and HOM2 are always more stable than the MO circuit for finite values of the dimer dissociation constant K_1, while HOM1 circuit shows the opposite. Furthermore, at tight repression ($s \to 0$), HET and HOM2 reach ultrastability even at moderate levels of expression efficiency.

Contrary to HET, where the protein-protein binding acts to deplete the minor protein and effectively suppresses the potential of state inversion, the stability of HOM2 is mainly driven by "inertia": Since the copy number of a protein monomer increases as β in steady state, the dimer copy number increases as β^2/K_1, leading to a disproportionately large number in the limit of strong-binding affinity and high expression efficiency. At $K_1 = 10$ nM and $\beta = 64$, the average copy number of the majority dimer is approximately 1788, in stark contrast with that of the minority dimer (approximately 3; see the supplemental files [8]), making a state inversion virtually impossible. In particular, note that the HOM2 switching rate decreases by 8 orders of magnitude when β increases from 2 to 16.

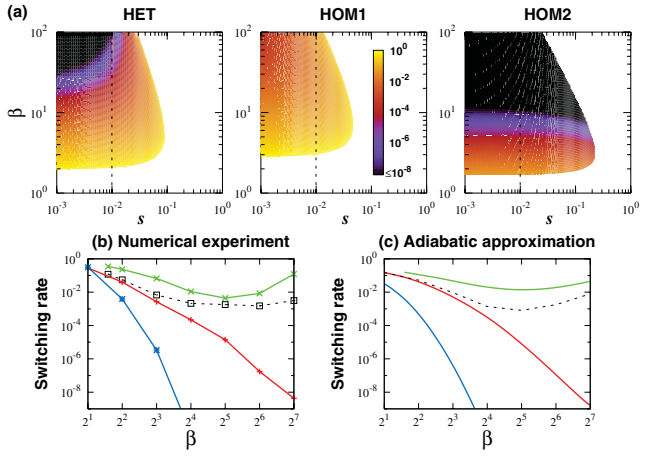

FIG. 4 (color). (a) Heat map of the inverse first-passage time in an adiabatic approximation. (b) Switching rates tallied from the exact time series. (c) An excerpt from the heat map along the trajectory, where $K_1 = 10$ nM and $s = 0.01$ (dashed lines in the heat maps). Color code: red, HET; green, HOM1; blue, HOM2; broken lines, MO. All the rates are scaled by the inverse lifetime of the protein monomer.

Discussion.—We have studied the tunability and stability, hence the potential for epigenetic memory, in three different realizations of a simple genetic toggle switch with the presence of protein dimers. We found that HET and HOM2 circuits access a much wider range of switching rates than MO circuits: HET and HOM2 circuits are always more stable than MO circuit, revealing ultrastability in a biologically relevant parameter range ($s < 0.01$, $10 < \beta < 100$) whereas HOM1 is even less stable and less tunable than MO. A large number of evolved genetic circuits are known to have dimeric transcription factors. Best known examples include the CI, TrpR, and leucine zipper proteins, which correspond to the HOM2 circuit class. This fact suggests that their tunable stability may be an outcome of natural selection on functional gene modules. Though the HET circuit, to our knowledge, is not found in naturally evolved systems, we suggest that this circuit design may find use in synthetic biology as part of engineered memory devices. HET may be a particularly well suited starting point for situations where there is a need for tight control of the number difference in protein populations.

The circuit parameters s (promoter leakage) and β (expression efficiency) are readily accessible to experimental manipulation. For example, to modulate the promoter leakage s, plasmid integration techniques can be used in a combinatorial manner [17] to vary the strength of steric hindrance by repressor proteins. The modulation of β is achieved by changing the synthesis rates of mRNA or protein and/or the degradation rate thereof. One possible way to implement the β tuning is to fuse the ssrA tag to the $3'$ end of each gene [18] so that the protein degradation is exposed to active proteolysis. A recent study shows that the use of active proteolysis significantly enlarges the parameter space for oscillation and thus improves the robustness of synthetic circuits under parametric uncertainty [19]. Finally, the dimer dissociation constant K_1 may also be modulated, as it is possible to either change the binding domain sequence of the repressor proteins or to treat the circuit with a small molecule that induces allosteric deformation in the binding site. Although many technological hurdles are still present for the realization of plug-and-play synthetic biology [20], the inherent tunability of the proposed gene switches makes them an ideal part for scalable functional modules.

This work was performed under the auspices of the U.S. DOE by LLNL under Contract No. DE-AC52-07NA27344 and funded by LDRD program Grant No. 06-ERD-061.

*eivind.almaas@ntnu.no

[1] L. Hartwell, J. Hopfield, S. Leibler, and A. Murray, Nature (London) **402**, C47 (1999).
[2] A. Barabási and Z. Oltvai, Nat. Rev. Genet. **5**, 101 (2004).
[3] J. Lisman, Proc. Natl. Acad. Sci. U.S.A. **82**, 3055 (1985); W. Xiong, Nature (London) **426**, 460 (2003).
[4] D. Rozanov, R. D'Ari, and S. P. Sineoky, J. Bacteriol. **180**, 6306 (1998); J. Little, D. Shepley, and D. Wert, EMBO J. **18**, 4299 (1999).
[5] M. Acar, J. Mettetal, and A. van Oudenaarden, Nat. Genet. **40**, 471 (2008).
[6] T. Gardner, C. Cantor, and J. Collins, Nature (London) **403**, 339 (2000).
[7] A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633 (1998); D. K. Hawley, A. D. Johnson, and W. R. McClure, J. Biol. Chem. **260**, 8618 (1985); M. Shea and G. Ackers, J. Mol. Biol. **181**, 211 (1985).
[8] See EPAPS Document No. E-PRLTAO-103-070928 for supplemental material. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[9] P. H. von Hippel and O. G. Berg, J. Biol. Chem. **264**, 675 (1989).
[10] A. Lipstat, A. Loinger, N. Q. Balaban, and O. Biham, Phys. Rev. Lett. **96**, 188101 (2006).
[11] M. A. Gibson and J. Bruck, J. Phys. Chem. A **104**, 1876 (2000).
[12] M. N. Artyomov, J. Das, M. Kardar, and A. K. Chakraborty, Proc. Natl. Acad. Sci. U.S.A. **104**, 18 958 (2007).
[13] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **94**, 018104 (2005).
[14] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).
[15] A. Raj *et al.*, PLoS Biol. **4**, e309 (2006).
[16] D. Roma, R. O'Flanagan, A. Ruckenstein, A. Sengupta, and R. Mukhopadhyay, Phys. Rev. E **71**, 011902 (2005).
[17] K. Murphy, G. Balázsi, and J. Collins, Proc. Natl. Acad. Sci. U.S.A. **104**, 12 726 (2007).
[18] C. Grilly, J. Stricker, W. L. Pang, M. R. Bennett, and J. Hasty, Mol. Syst. Biol. **3**, 127 (2007).
[19] W. W. Wong, T. Y. Tsai, and J. C. Liao, Mol. Syst. Biol. **3**, 130 (2007).
[20] D. Endy, Nature (London) **438**, 449 (2005).