LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-228269

Modeling the Mechano-Chemistry of NTPases

Jianhua Xing

February 23, 2007

Disclaimer

Auspices Statement

This work was performed under the auspices of the U. S. Department of Energy (DOE) by the University of California, Lawrence Livermore National Laboratory (LLNL) under Contract No. W-7405-Eng-48. The project (*06-ERI-0004*) was funded by the Laboratory Directed Research and Development Program at LLNL.

FY06 LDRD Final Report
Modeling the Mechano-Chemistry of NTPases
LDRD Project Tracking Code: 06-ERI-004
Jianhua Xing, Principal Investigator

Abstract

This project is to develop theoretical framework for protein motors based on experimental data. Protein motors use chemical and electrochemical energies to perform mechanical work. Protein motors are machines of life. They are essential for many biological processes, including cell division, DNA transcription, replication, etc. Understanding the working mechanisms of protein motors has both scientific and medical/clinical significances, including revealing the physiological origins of certain diseases, designing of drugs against pathogens. Experiments with new techniques, especially recent advances in single molecule force measurements, have accumulated a large amount of experimental data that requires systematic theoretical analysis.
We worked out a theoretical analysis on protein fluctuations to explain the recent single molecule experiment on dynamic disorders, proposed a new mechanism to explain mechanical signal propagation through the allosteric effect, a fundamental property of proteins, and examined the dynamic disorder effects on protein interaction networks. We also examined various theoretical formulations describing mechanical stress propagation in proteins, and derived mathematical formula for various approximate methods solving the mathematical equations.

Introduction / Background

Protein motors are essential for almost every biological process. Studying protein motors has significance in both basic science and medical applications (see below for details). The results will be disseminated to broader audience through publication and scientific conferences. We will also make the generated computer codes freely accessible to the public.

Biological/medical importance of protein motors

Protein motors are proteins which can perform mechanical work [1]. They are machines made by nature that share some common features with macroscopic manufactured machines, such as the use of repeating motor cycles. There are also significant differences. Due to nano-scale sizes, motions of protein motors are dominated by thermal fluctuations. Protein motors are driven by either transmembrane ion-motive force (a trans-membrane electrical potential and ion concentration gradient) or through release of chemical energy (mainly hydrolysis of triphosphates such as ATP and GTP). Almost every biological process requires protein motors, including DNA transcription, replication, cell division, muscle contraction—even ATP synthesis.

Not surprisingly, many protein motors play an essential role in some pathologies. For example, HCV NS3, a helicase that unwinds DNA for replication, is essential for Hepatitis C virus growth [2, 3]. The protein FtsK coordinates bacteria proliferation by translocating DNA[4]. Viruses, such as herpes virus and adenovirus, use DNA and RNA packaging motors to pack their genome into preformed protein capsids [5-8]. PilT and PilB are two protein motors used by many Gram-negative bacteria to mediate adhesion of type

IV pili to host mammalian cells [9, 10]. The motility of many bacteria, including many pathogens, depends on a complex of proteins known as the bacteria flagellar motor (BFM). Given this wide variety of essential motor protein functions, it is not surprising that many motor proteins are drug targets against various diseases. Understanding their functional mechanism provides guidance for developing drugs that block or impair the normal functions. Also promising is the engineering of some protein motors to perform curing functions. For example, the virus packaging motor may be used for delivery of therapeutic genes[8]. Extensive efforts have been made to develop micron scale machines to perform precise localized surgeries. At such small sizes, fluid transportation is dominated by low Reynold number dynamics (i.e., motion within a highly viscous medium, like a ball in molasses), and becomes highly inefficient if driven by pressure gradient. The bacterial flagellar motor could be used to pump fluids in these devices.

Recent advances in single-molecule techniques have allowed precise measurements of the mechano-chemical behaviors of single protein motor in assays. However, in cells more than one motor proteins may attach to a vesicle. The mechano-chemical behaviors of the system are not a simple summation of single motor behaviors. Much less experimental investigations have been conducted on vesicle transport *in vivo*[11-13]. Our collaborators at Wake Forest use motion enhanced differential interference contrast microscopy to track small vesicles and thus obtain their transporation rates in neurites[11]. Since experiments *in vivo* is less controlled than those *in vitro*, theoretical modeling is necessary to analyze the data.

Mathematical modeling of protein motors

Mathematical modeling serves to integrate and transform quantitative and qualitative experimental information into coherent and consistent physical models that yield quantitative, experimentally falsifiable predictions. More importantly, modeling can suggest new experiments and new quantities to measure.

Modeling of protein motors can be classified into three general approaches. (i) Molecular dynamics (MD) simulations account for the motion of every atom, and sometimes the surrounding water and other chemical species, but have serious limitations: the size of a protein motor and the relevant time scale (in milliseconds) are usually beyond current computer capacity, and the required atomic structures are not available for most motor proteins. (ii) Kinetic (rate-equation type) models represent the transitions between a small number of chemical (Markov) states. (iii) Markov-Fokker-Planck (MFP) models lie somewhere in between the atomic detail of MD and the phenomenology of kinetic models (4). These models describe the geometric motion along a few ‘collective coordinates’ that describe the protein’s major conformational movements. Motion is driven by forces derived from a set of potential functions assigned to each coordinate, with Markov jumps between the potentials corresponding to chemical transitions. MFP models generalize kinetic models by replacing the discrete kinetic states with potential functions representing collective spatial coordinates. Given that protein motors are machines, continuum MFP models provide more appropriate treatments on the mechanical aspects of protein motors than discrete kinetic models do. Therefore, the proposed research will focus on the coarse-grained MFP models. We emphasize that MFP models complement MD and kinetic models; each has its proper place in understanding a mechanochemical system as complex as a protein motor.

Our protein motor modeling philosophy coincides with the systems approach receiving great attention recently in biology. In particular, each protein motor can be viewed as a miniature system. While each experimental technique (including atomistic level simulation) focuses on one or a few aspects of the system, coarse-grained modeling serves as the unique role of integrating all the information into a coherent, consistent, and global physical model of the system. Consequences and functions of some

observations (e.g., mutations of certain amino acid residues) can only be fully appreciated in the context of overall motor dynamics.

Continuum potential description

Since the discrete rate equation formalism is widely used in describing biochemical reactions, here we briefly discuss its relations to more detailed continuum description.

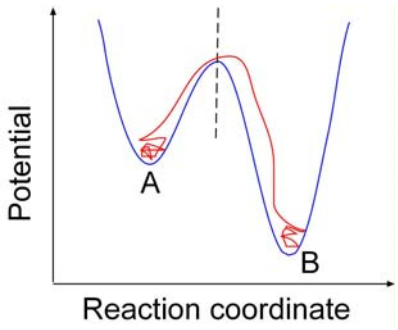


Figure 1 Free energy profile (blue) for a typical chemical reaction.

Ultimately all the molecular interactions can be described by continuum potentials. For a typical free energy profile shown in Figure 1, the system spends most of time wandering inside the potential wells, while transitions between the two regions are nearly instantaneous. The system can be "coarse-grained" into two discrete states (A and B in Figure 1), and transitions can be described by the familiar rate equations. Calculation of the corresponding rate constants from underlying continuum potentials is a major research field in chemical physics[14].

However, since rate equation expressions are approximation of continuum dynamics, some details are inevitably lost. While its level of details are satisfactory on describing normal enzymatic reactions, rate expressions are oversimplified for describing protein motors, because one wants to know not just the initial and final states, but also the transition process as well. For protein motors, the intermediate process may not be fast enough to be treated as instantaneous jumps, and this limitation becomes more apparent as current single molecule techniques provide details not accessible by bulk measurement techniques.

Continuum potential descriptions have been receiving increasing recognition by both theoreticians and experimentalists in the protein motor community [15-17]. To further illustrate the advantage of this approach, an analogy to an internal combustion engine may be helpful. While the engine cycle can be divided into states that include intake, compression, combustion, and exhaust, this discrete description does not yield a mechanistic understanding of how these states come about. Rather, one wants to know the physical design of the engine and how each moving part coordinates with the others. The same is true for our understanding of protein motors with the important difference that the dynamics of protein motors is not deterministic and requires consideration of thermal fluctuations.

Research Activities and Results

My computer plan was not approved until May 10th, 2006. Therefore the research has been focused on theoretical analysis and numerical calculations with minimal computational requirement.

Protein fluctuations

Protein motors function through substrate-binding regulated conformational changes. We used Langevin dynamics formulism to describe conformational motions. The Xie group at Harvard use single molecule experimental technique showed that protein fluctuations show long-time memory effect. It is controversial on how to understand the results. With Dr Ken Kim, I worked out a theoretical explanation of their experiment. With two fitting parameters which agree well with other independent studies, our theoretical results agree well with the experimental data from the Xie group. The paper was published in Phys. Rev. E. 74:061911 (2006), and was selected to publish at the APS Virtual Journal of Biological Physics (Vol. 13, Issue 1, 2007).

Allosteric effect

Allosteric effect refers to the property that activity (substrate binding, reaction, etc) at one site of the protein can affect the behavior of another remote site of the protein. It is a fundamental property of many proteins, and is closely related to functioning mechanisms of protein motors. The latter depends on ligand binding to affect reactivity at another site. Classical textbook mechanisms focus on thermodynamic behaviors of the protein under ligand binding. The above work on protein fluctuations inspired me to propose a different mechanism of allosteric regulation. The attached manuscript is ready for submission.

Dynamic disorders and protein interaction network

One usually assumes that a chemical reaction is described by a rate constant. This is the basic framework for describing biochemical reactions and biological networks. The experiments by the Xie group showed that an enzymatic reaction rate is actually not a constant. This observation may have profound impact on our understanding of dynamics of biological systems. In collaboration with several other researchers, I analyzed the effect of dynamic disorders on a small protein interaction network, the classical Goldbeter-Koshland model. The attached manuscript is about ready for submission.

Bacterial flagellar motor

I instructed an Oxford student (Fan Bai in Richard Berry's group) to perform studies on the flagellar motor. Our original work on the motor is a phenomenological model. In this work we added detailed information based on recent experimental studies, and made predictions to compare with single molecule data obtained in the Berry group. For the first half year, I helped the student to develop computer codes to perform Fokker-Planck and Langevin dynamics calculations. Currently I am instructing the student to collect computational results. Some results are shown in Figure 2 and Figure 3. A manuscript is in preparation.

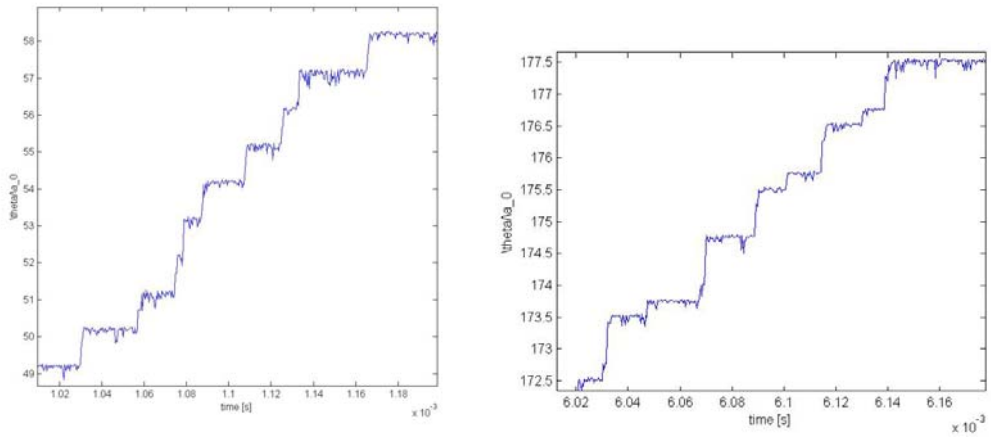


Figure 2 Our model predicts that the motor is a stepper with a stepsize $2\pi/26$ with one stator (left). With two stators the stepsizes depend on relative arrangements of the stators (right). The predictions are consistent with single molecule data[18].

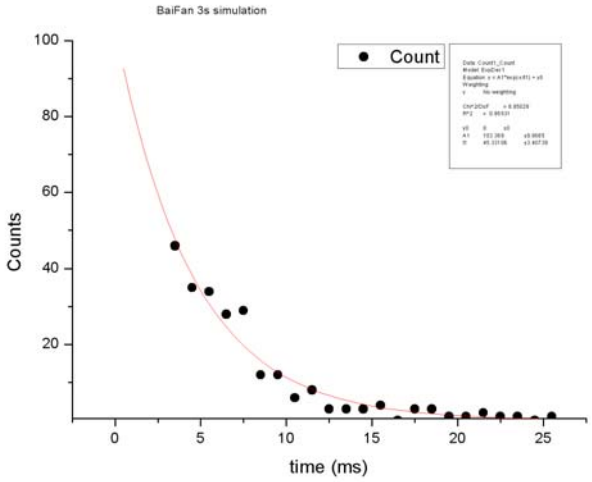


Figure 3 Dwelling time distributions.

Theoretical modeling framework

Meanwhile, I have been working on a manuscript describing the general theoretical modeling framework. I derived mathematical formula for various approximation methods. Numerical tests of these methods are underway.

With Dr. Ken Kim, I studies various ways of theoretical expressions to describe mechanical stress propagation. Currently we are collaborating with a Virginia Tech group for tensor visualization. Application on the clamp/clamp loader system (in collaboration with Dr. Daniel Barsky at LLNL) is underway.

I also co-instructed a Berkeley student (Joshua Adelman in Professor George Oster's group) to develop a general computer code and model DNA packaging motor. The code has been tested to work.

Summary

This project received about half-year of support. During this period, we managed to get one paper published, and several manuscripts in preparation.

Acknowledgements

Given the unexpected difficulties with my computer plan, my host Dr Chris Orme provided invaluable help on my adjustment, and on every issue raised up with my research and staying at the laboratory. I also received continuous advices and help from Drs. Ken Kim, Daniel Barsky, and Mike Surf.

References

1 **J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer, Sunderland, MA, 2001).**

2 **S. Dumont, W. Cheng, V. Serebrov, et al., Biophys. J. 88, 350a (2005).**

3 **V. Serebrov and A. M. Pyle, Nature 430, 476 (2004).**

4 **H. Capiaux, C. Lesterlin, K. Perals, et al., EMBO Reports 3, 532 (2002).**

5 **W. C. Earnshaw and S. R. Casjens, Cell 21, 319 (1980).**

6 **Y. R. Chemla, K. Aathavan, J. Michaelis, et al., Cell 122, 683 (2005).**

7 **L. W. Black, Annu. Rev. Microb. 43, 267 (1989).**

8 **P. Guo, Methods Mol Biol 300, 285 (2005).**

9 **B. Maier, I. Chen, D. Dubnau, et al., 11, 643 (2004).**

10 **A. J. Merz, M. So, and M. P. Sheetz, Nature 407, 98 (2000).**

11 **D. B. Hill, M. J. Plaza, K. Bonin, et al., Eur. Biophys. J. Biophys. Lett. 33, 623 (2004).**

12 **R. D. Allen, J. Metuzals, I. Tasaki, et al., Science 218, 1127 (1982).**

13 **S. T. Brady, R. J. Lasek, and R. D. Allen, Science 218, 1129 (1982).**

14 **P. Hanggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. 62, 254 (1990).**

15 **D. M. Goedecke and T. C. Elston, J. Theor. Biol. 232, 27 (2005).**

16 **H. Itoh, A. Takahashi, K. Adachi, et al., Nature 427, 465 (2004).**

17 **D. Keller and C. Bustamante, Biophys. J. 78, 541 (2000).**

18 **Y. Sowa, A. D. Rowe, M. C. Leake, et al., Nature 437, 916 (2005).**

Protein fluctuations and breakdown of time-scale separation in rate theories

Jianhua Xing[1,2,*] and K. S. Kim[1]
[1]*Chemistry and Material Science Directorate, University of California & Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[2]*Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0406, USA*
(Received 10 May 2006; revised manuscript received 3 October 2006; published 27 December 2006)

A long-time fluctuation correlation function with a power-law form has been observed in recent single-molecule experiments by the Xie group. By analyzing the dynamics of an elastic network model (ENM) under white noise, we show that the observed long-time memory kernel can be explained by the discrepancy between the experimentally measured coordinate (or the coordinate directly coupled to protein function) and the minimum energy path of the system. Consequently, the dynamics of the measured collective coordinate has contributions from degrees of freedoms with a broad distribution of time scales. Our study also implies that the widely used ENM Hamiltonian should be viewed as a coarse-grained model of a protein over a rugged energy landscape. Large effective drag coefficients are needed to describe protein dynamics with the ENM's.

DOI: 10.1103/PhysRevE.74.061911 PACS number(s): 87.15.He, 87.15.Ya, 87.16.Ac, 82.37.−j

I. INTRODUCTION

Rate processes are ubiquitous in physics, chemistry, and biology. The development of reaction rate theories is a classical topic of theoretical physics and chemistry [1]. A basic physical picture is given in the seminal paper by Kramers [2] (see Ref. [1] for earlier and subsequent work). For a dynamical many-body system, a basic premise behind the construction of many reaction rate theories is the existence of one special degree of freedom (DOF) called the reaction coordinate (RC), whereby the trajectory of the species along this coordinate results in the chemical reaction. Mathematically it is normally the minimum-energy path (MEP) (or the intrinsic reaction coordinate) connecting the reactant and the product along the mass-weighted multidimensional potential energy surface (see Fig. 1) [3]. One usually assumes a separation of time scales between the dynamics along the MEP and along the remaining DOF orthogonal to the MEP. Consequently, the system dynamics can be well described by a Langevin-type or generalized Langevin-type dynamics with-short time memory kernels [4],

$$m\frac{d^2x}{dt^2} = -\frac{dU}{dx} - \zeta_g \int_0^t d\tau K(t-\tau)\frac{dx(\tau)}{d\tau} + f(t), \qquad (1)$$

where m is the reduced mass, x represents the coordinates of the MEP, U is the potential along MEP, ζ_g is the drag coefficient, K is the memory kernel which usually take the form of a Dirac δ function (for Langevin dynamics) or a fast decaying function of $(t-\tau)$ (e.g., an exponentially decaying function in the Grote-Hynes treatment [5]), and f is the fluctuation force. The generalized Langevin equation (1) can be derived formally using the Mori-Zwanzig projection operator formalism with temporal-spatial coarse graining [4,6,7]. The basic idea is that a system can be divided into two subspaces with the dynamics treated explicitly and implicitly, respectively. The projection formalism involves the construction of an operator which projects the full dynamics of system onto a subspace spanned by the explicit DOF. The projection formalism does not eliminate the implicit DOF, but preserves their influence or back reaction on the reduced dynamics through the appearance of a memory term and a stochastic forcing term. Under the condition that the dynamics of the explicit and implicit DOF be slow and fast modes, respectively, and that there exist a clear time-scale separation, the memory term decays quickly.

Proteins are flexible entities. Numerous experimental and theoretical studies have found that the structural fluctuations of proteins are strongly correlated with their function [8–12]. A widely used technique to study protein fluctuations is to calculate the potential of mean force (PMF) along the RC. On obtaining the PMF, one assumes that all the DOF or-

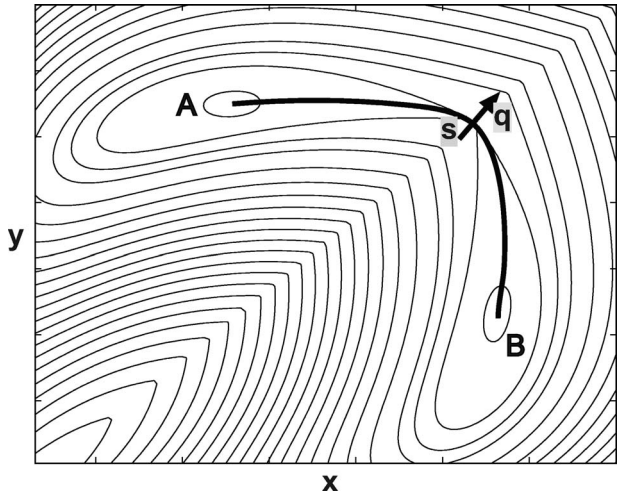


FIG. 1. Schematic illustration of a potential energy surface and the minimum-energy path (s). The shown 2D system can be described by either a coordinate system with the MEP and the orthogonal coordinate q or an x-y coordinate system. A projection along the MEP may result in a 1D generalized Langevin equation with short-time memory kernels. However, if the projection coordinate (e.g., x, set by either experiments or functional relevance) deviates from the MEP, a long-tail memory kernel may be expected.

*Corresponding author. Electronic address: xing3@llnl.gov

thogonal to the RC adjust to motion along the RC adiabatically. In other words, there is clear time-scale separation between dynamics along the RC and along the remaining DOF. Recent single-molecule measurements by the Xie group have shown a power-law memory kernel for the fluctuations within proteins of various systems described by a generalized Langevin equation [13,14]. The existence of a long-time memory effect implies the breakdown of the time-scale separation assumption, thus initiating several theoretical investigations to clarify the mechanism behind the power-law decay. Granek and Klafter explain the observation by using a fracton dynamics model [15]. Debnath and co-workers [16] and Tang and Marcus [17] show that distance fluctuations of one-dimensional polymers give the observed power-law decay of the memory kernel. Here, we propose a dynamic version of the widely used elastic network model which is simple and yet captures the essential physics. The experimentally measured coordinate (or the coordinate directly coupled to the protein function) may not coincide with the minimum-energy path of the system; e.g., consider the projection is along x in Fig. 1. Then the above-mentioned time-scale separation condition is not satisfied and a long-time memory term will be expected.

II. THEORY AND NUMERICAL RESULTS

In the present work, we adopted the widely used elastic network model (ENM) to analyze the experiment by Min *et al.* [14]. The ENM is a coarse-grained model and gives a reasonable description of protein fluctuations [18–20]. An ENM represents a protein by a network of elastically coupled N nodes (usually the C_α atom positions), $\mathbf{q}=(\mathbf{q}_1,\ldots,\mathbf{q}_N)$, with the following simple interaction form:

$$V=\frac{1}{2}c\sum_{i\neq j}h(r_{cut}-r_{ij}^e)(r_{ij}-r_{ij}^e)^2, \tag{2}$$

where the superscript e represents the equilibrium structure, c is a universal spring constant, h is a Heaviside function, r_{cut} is a cutoff distance, and $r_{ij}=|\mathbf{q}_i-\mathbf{q}_j|$ and $r_{ij}^e=|\mathbf{q}_i^e-\mathbf{q}_j^e|$.

The above potential form can be approximated by expanding to quadratic terms. By diagonalizing the Hessian, one can transform to normal-mode coordinates with $\Delta\mathbf{q}\equiv\mathbf{q}-\mathbf{q}^e=\mathbf{\Gamma Q}$, where $\mathbf{\Gamma}$ is composed of the eigenvectors of the Hessian. To study the dynamics, we include the effects of solvent and coarse-grained degrees of freedom in the ENM as dissipation terms, so the equations of motion are given by a set of overdamped Langevin equations (see also [21–23])

$$-\lambda_{i\alpha}Q_{i\alpha}-\zeta\frac{dQ_{i\alpha}}{dt}+f(t)=0, \tag{3}$$

where $Q_{i\alpha}$ and $\lambda_{i\alpha}$ are the $[(i-1)\times 3+\alpha]$th normal-mode coordinate and the corresponding eigenvalue. Here, for convenience of discussion, we use two indices to label the normal mode, with $i=1,\ldots,N$ and $\alpha=1$, 2, 3. From Eq. (3), the coordinate autocorrelation function is given by

$$C_Q^{i\alpha}(t)\equiv\langle Q_{i\alpha}(t)Q_{i\alpha}(0)\rangle=\frac{k_BT}{\lambda_{i\alpha}}\exp\left(-\frac{\lambda_{i\alpha}}{\zeta}t\right), \tag{4}$$

with k_B the Boltzmann constant and T the temperature. The Laplace transform is

$$\tilde{C}_Q^{i\alpha}(s)=\frac{k_BT}{\lambda_{i\alpha}}\frac{1}{s+\lambda_{i\alpha}/\zeta}. \tag{5}$$

To compare with the experimental results of Min *et al.* [14], one needs to calculate the autocorrelation function of the distance between two ENM nodes, which is approximately given by

$$\begin{aligned}C_r(t)&\equiv\langle(r_{ij}(t)-r_{ij}^e)(r_{ij}(0)-r_{ij}^e)\rangle\\&\approx\left(\frac{1}{r_{ij}^e}\right)^2\langle(\mathbf{q}_i^e-\mathbf{q}_j^e)\cdot[\Delta\mathbf{q}_i(t)-\Delta\mathbf{q}_j(t)]\\&\quad\times(\mathbf{q}_i^e-\mathbf{q}_j^e)\cdot[\Delta\mathbf{q}_i(0)-\Delta\mathbf{q}_j(0)]\rangle\\&=\sum_{k\beta}\left\{\left(\frac{1}{r_{ij}^e}\right)^2\sum_\alpha(q_{i\alpha}^e-q_{j\alpha}^e)^2[(\bar{\Gamma}_{i\alpha,k\beta})^2+(\bar{\Gamma}_{j\alpha,k\beta})^2]\right.\\&\quad\left.\times\langle Q_{k\beta}(0)Q_{k\beta}(0)\rangle\right\}\frac{\langle Q_{k\beta}(0)Q_{k\beta}(t)\rangle}{\langle Q_{k\beta}(0)Q_{k\beta}(0)\rangle}\\&=\sum_{k\beta}p_{k\beta}\frac{C_Q^{k\beta}(t)}{C_Q^{k\beta}(0)}.\end{aligned} \tag{6}$$

In the above expression, the matrix $\bar{\Gamma}$ is the reduced Γ after the elimination of the zero-frequency modes corresponding to the three translational degrees of freedom of the center of mass and three rotations. The normal-mode correlation functions $C_Q^{k\beta}(t)$ are given by Eq. (4). Equation (6) reveals that the distance correlation function, which is given by a linear combination of the single-exponential decaying normal-mode correlation functions, may show multiexponential decay. If the generalized Langevin equation (1) with a harmonic potential is used to model the residue-residue distance fluctuation, the memory kernel in Laplace space is given by [14]

$$\tilde{K}(s)=\frac{m\omega^2}{\zeta_g}\frac{\tilde{C}_r(s)}{C_r(0)-s\tilde{C}_r(s)}, \tag{7}$$

where $\tilde{C}_r(s)$ is the Laplace transform of the distance correlation function $C_r(t)$,

$$\tilde{C}_r(s)\equiv\sum_{k\beta}p_{k\beta}\frac{\tilde{C}_Q^{k\beta}(s)}{C_Q^{k\beta}(0)}. \tag{8}$$

Equations (6)–(8) are the central results of this work. They show that a long-time memory kernel for the distance fluctuation can exist even if all the normal modes are described by the Langevin dynamics with δ memory kernels. The reason is that the distance fluctuation coordinate is not along any of the normal-mode coordinates and has contributions from a large number of normal modes [see Fig. 3(b) below].

We apply the above analysis here to the protein complex formed between fluorescein (FL) and monoclonal antifluo-

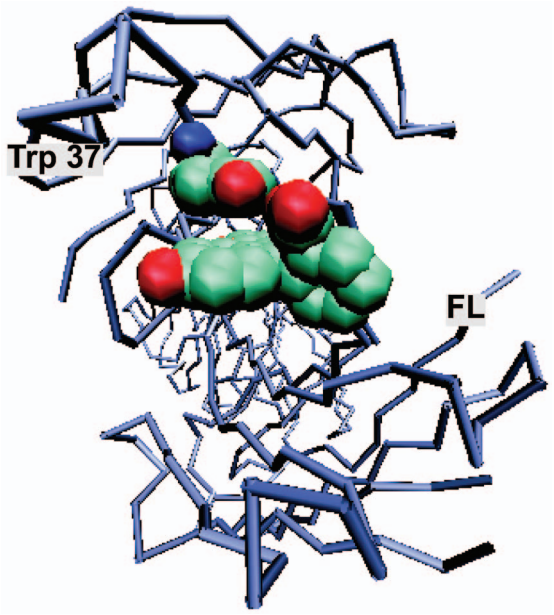


FIG. 2. (Color) Structure of the FL–anti-FL protein complex. In ENM, the protein structure 1FLR of anti-FL is modeled by 437 nodes representing the Ca atoms with an extra node representing the center of the aromatic ring of Trp37 and the fluorecein is modeled by nodes locating at the centers of the three aromatic rings. The distance fluctuation between the center of Trp37 ring and the FL ring formed by C3-C8 is calculated to compare with the experiment.

rescein 4-4-20 (anti-FL) (PDB label 1FLR; see Fig. 2). The spring constant c in Eq. (2) is determined by fitting to $C_r(0)$ and the drag coefficient ζ by fitting the experimental data of the distance correlation function [14]. In the calculations, the temperature is assumed as $T=298$ K and a cutoff distance $r_{cut}=10$ Å used in the literature is adopted [19], $c=1.4$ kcal/(mol/Å^2), comparable to the value 1.0 ± 0.5 kcal/(mol Å^2) used in the ENM literature [19], $\zeta=1k_BT$ s/Å^2. Figure 3(a) compares the calculated distance correlation function using Eq. (6) and the experimental data of Min *et al.* Given the simple form of the ENM, the agreement is remarkable. At large t, there is larger discrepancy between the theoretical result and the experimental data in Fig. 3(a). This discrepancy implies that contributions from some low-frequency modes are underestimated with the current model.[1] Models with different cutoff distances also give reasonable fittings. Figure 3(b) shows the normalized contribution of each normal mode given by Eq. (6). While several low-frequency modes make significant contributions, those from other modes cannot be neglected due to the large number of degrees of freedom involved. Figure 3(c) shows the Laplace transform of the memory kernel calculated by Eq. (7), which has approximately a power-law form. Figure 3(c) also shows that distance correlations between the FL and several other residues have similar behaviors, which was also

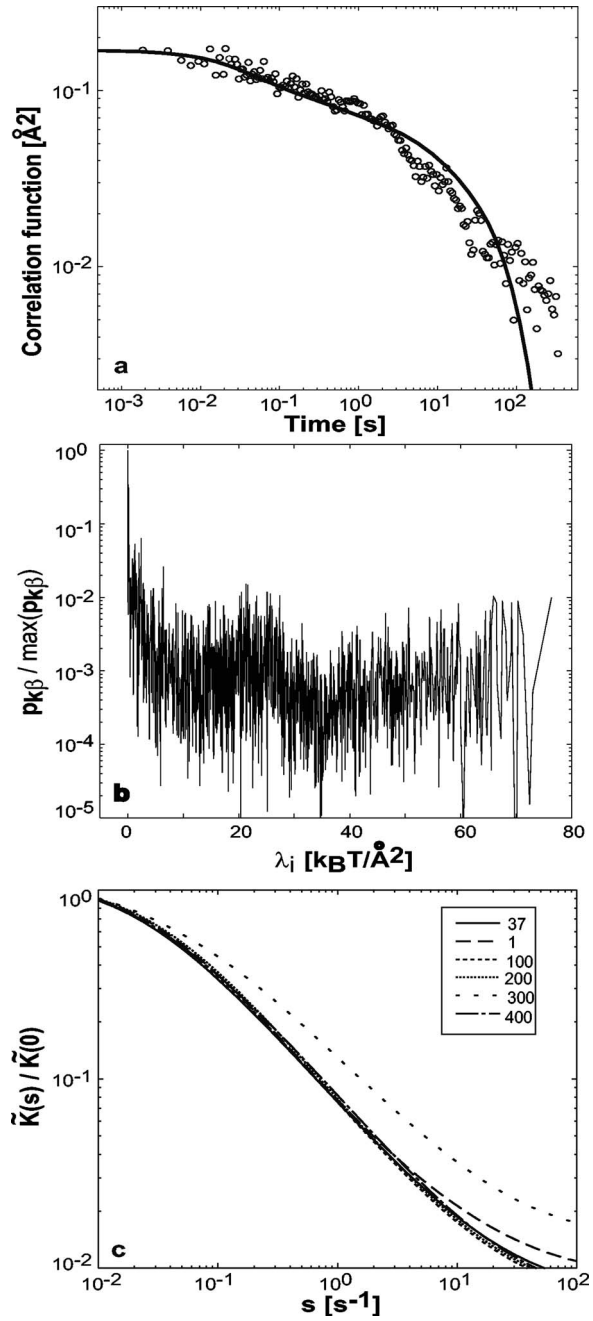


FIG. 3. The distance correlation function between Tryp 37 and the FL. (a) Comparison of the calculated distance correlation function with Eq. (6) (solid line) and the experimental data by Min *et al.* [14] (circles). (b) The normalized contribution of each normal mode given by Eq. (6). (c) Laplace transform of the memory kernel calculated by Eq. (7). Here the calculated memory kernels from the distance correlation functions between the FL and residue 37 as well as residues 1, 100, 200, 300, and 400 are shown.

observed in the one-dimensional polymer study of Tang and Marcus [17]. This implies that the lack of time-scale separation is a general phenomenon in protein dynamics.

III. DISCUSSION

The phenomenon of dynamic disorder, or rate constant fluctuations, has been widely studied [24–27]. The experi-

[1]To focus on the essential physics, we adopted the simplest version of the ENM. Improvement of the fitting is expected with some refined but more complex models discussed in the literature.

mental observations of the Xie group further demonstrate dynamic disorder at single-molecule levels. The experiments reveal that caution should be taken on applying rate theories to complex biological systems. The legitimacy of the underlying assumptions of a rate theory should be reexamined. The reaction coordinate framework discussed in the Introduction can break down in two respects. First, a reaction coordinate with slow dynamics may not exist if the system involves broad and continuous time scales. Second, even if a slow reaction coordinate (the MEP) can be defined, the coordinate relevant to experimental measurements or to protein functions may not coincide with the MEP, and this discrepancy may result in the breakdown of the time-scale separation between the tagged coordinate and the remaining DOF. Consequently, to describe the system dynamics, a long-time memory term may be necessary or extra DOF need to be treated explicitly. These observations may have significant relevance in any study involving the dynamics of biological macromolecules. Here we want to mention a similar situation in protein motor studies. Protein motors use chemical and electrochemical energies to perform mechanical work and are essential for many biological processes [28–32]. A protein motor is described by chemical reaction coordinates and mechanical motion coordinates which are coupled together, analogous to the electron-transfer coordinate and the protein fluctuation coordinate in the FL–anti-FL system. There is usually no time-scale separation between the different degrees of freedom. Min *et al.* related the observed long-time memory kernel to the observation that there may be no well-defined single-valued rate constant for an enzymatic reaction [33,34]. Studies show that theoretical treatments beyond discrete rate equations are necessary to understand some mechanochemical properties of a protein motor [35–37]. For example, a distribution of rate constants is essential to explain the long-standing puzzle of the motor torque-speed relationship of the bacterial flagellar motor [38,39] [comparing Eq. (6) of Ref. [39] and Eq. (5) of Ref. [33]]. Further studies are needed to examine the implications of the experimental observations of Min *et al.* to the understanding of other biological systems.

To fit the experimental data, we used a drag coefficient $\zeta = 1 k_B T$ s/Å^2. This value is orders of magnitude higher than the typical drag coefficient of a polymer [17]. This discrepancy may call into question the validity of the distance fluctuation models as discussed in the present work and in the work of Tang and co-workers [17,40]. However, the discrepancy can be reconciled by the fact that the elastic network model is a coarse-grained model. Many experimental and theoretical studies (especially in the protein folding community) show that a system complex such as a protein possesses a rugged energy [41,42]. The ruggedness of the energy landscape may lead to the conclusion that a normal-mode analysis cannot work since it only characterizes the potential near one local minimum. On the other hand, the smooth potential used by an elastic network model should be understood as an effective potential after averaging out the local rugged fluctuations [43,44]. That explains why the elastic network models are surprisingly successful on describing large-scale protein fluctuations (notice that the short-range fluctuations of the potentials have a marginal effect on the equilibrium properties such as the B factor at experimental resolutions) [19]. Consequently, the effective drag coefficients used in a dynamic version of the elastic network model are different from the bare drag coefficients and the difference can be large. To further illustrate this, we refer to the work of Zwanzig [45]. Zwanzig proposed a model describing diffusion in a rugged potential and derived an expression for the effective diffusion coefficient. His results show that the ruggedness of the potential can dramatically reduce the diffusion coefficient at low temperatures. Zwanzig looked at diffusion on a length scale much larger than the ruggedness and, in effect, replaced the original rugged potential with an effective smooth potential, integrating out the rapid small fluctuations. Carrying this analogy to our case, the effective drag coefficients used for an elastic network model should be normal-mode frequency dependent (see Fig. 4). For high-frequency modes with a length scale comparable to the characteristic length scale of the rugged potential, the values should approach the bare drag coefficients. For low-frequency modes with a length scale much larger than the characteristic length scale of the rugged potential, the drag coefficients reach renormalized values according to the analysis of Zwanzig. A set of normal-mode frequency-dependent drag coefficients may help fit the experimental data. Our current treatment with a single value of the drag coefficient is oversimplified and requires further study. Another prediction of the present model is that the effective drag coefficients and the distance autocorrelation function are expected to be highly temperature dependent. The experiment setup by the Xie group may serve as a tool to detect ruggedness of the protein energy landscape (see also Ref. [46]).

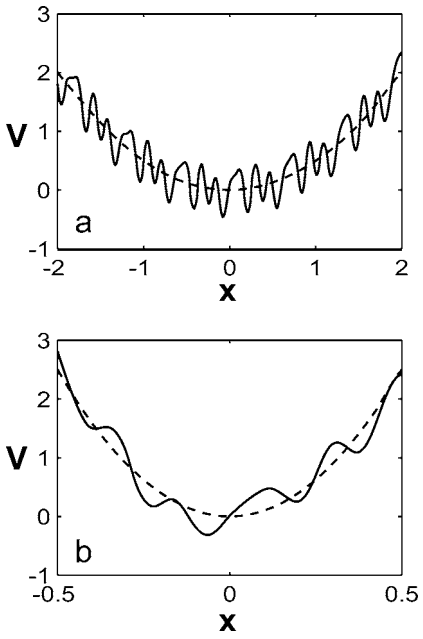


FIG. 4. Schematic illustration of rugged energy landscapes (solid lines). Also shown are renormalized smooth potentials with high (a) and low (b) frequencies (dashed lines). The units of distance and energy are arbitrary and only for illustration purposes. The existence of potential roughness may have different effects on the effective diffusion along the potential landscapes in the two cases.

ACKNOWLEDGMENTS

We would like to thank Wei Min (Harvard University) for providing us his data and pointing to Ref. [17], Professor Robert Jernigan and Dr. Andrzej Kloczkowski (Iowa State University) for pointing to Refs. [21–23], and Dr. Wenjun Zheng (NIH), Dr. Daniel Barsky (LLNL), and Dr. Michael Surh (LLNL) for many fruitful discussions. J.X. is supported by a Lawrence Livermore National Laboratory Directed Research and Development grant. K.K. would like to acknowledge the support by the DOE/The University of California Merced Center for a Computational Biology Grant No. DE-FG01–04ER04–15. This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory, under Contract No. W-7405-Eng-48.

[1] P. Hanggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[2] H. Kramers, Physica (Utrecht) **7**, 284 (1940).
[3] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
[4] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).
[5] R. F. Grote and J. T. Hynes, J. Chem. Phys. **73**, 2715 (1980).
[6] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).
[7] R. Zwanzig, in *Lectures in Theoretical Physics*, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience, New York, 1961), p. 106.
[8] S. H. Northrup *et al.*, Proc. Natl. Acad. Sci. U.S.A. **79**, 4035 (1982).
[9] M. F. Perutz and F. S. Mathews, J. Mol. Biol. **21**, 199 (1966).
[10] H. Frauenfelder, G. A. Petsko, and D. Tsernoglou, Nature (London) **280**, 558 (1979).
[11] T. W. Allen, O. S. Andersen, and B. Roux, J. Gen. Physiol. **124**, 679 (2004).
[12] M. Karplus and G. A. Petsko, Nature (London) **347**, 631 (1990).
[13] W. Min *et al.*, Acc. Chem. Res. **38**, 923 (2005).
[14] W. Min *et al.*, Phys. Rev. Lett. **94**, 198302 (2005).
[15] R. Granek and J. Klafter, Phys. Rev. Lett. **95**, 098106(1) (2005).
[16] P. Debnath *et al.*, J. Chem. Phys. **123**, 204903 (2005).
[17] J. Tang and R. A. Marcus, Phys. Rev. E **73**, 022102 (2006).
[18] M. M. Tirion, Phys. Rev. Lett. **77**, 1905 (1996).
[19] A. R. Atilgan *et al.*, Biophys. J. **80**, 505 (2001).
[20] W. J. Zheng and S. Doniach, Proc. Natl. Acad. Sci. U.S.A. **100**, 13253 (2003).
[21] A. Erkip and B. Erman, Polymer **45**, 641 (2004).
[22] A. Kloczkowski, J. E. Mark, and H. L. Frisch, Macromolecules **23**, 3481 (1990).
[23] W. W. Graessley, Macromolecules **13**, 372 (1980).
[24] R. Zwanzig, Acc. Chem. Res. **23**, 148 (1990).
[25] H. Frauenfelder, P. G. Wolynes, and R. H. Austin, Rev. Mod. Phys. **71**, S419 (1999).
[26] M. Karplus, J. Phys. Chem. B **104**, 11 (2000).
[27] X. S. Xie, J. Chem. Phys. **117**, 11024 (2002).
[28] G. Banting and S. Higgins, *Molecular Motors* (Portland Press, London, 2000).
[29] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer, Sunderland, MA, 2001).
[30] R. D. Vale, J. Cell Biol. **150**, 13 (2000).
[31] T. Ogura and A. J. Wilkinson, Genes Cells **6**, 575 (2001).
[32] J. Ye *et al.*, Biochim. Biophys. Acta **1659**, 1 (2004).
[33] B. P. English *et al.*, Nat. Chem. Biol. **2**, 87 (2006).
[34] W. Min and X. S. Xie, Phys. Rev. E **73**, 010902 (2006).
[35] P. Reimann, Phys. Rep. **361**, 57 (2002).
[36] J. Xing, H.-Y. Wang, and G. Oster, Biophys. J. **89**, 1551 (2005).
[37] F. Julicher, A. Ajdari, and J. Prost, Rev. Mod. Phys. **69**, 1269 (1997).
[38] H. C. Berg, Annu. Rev. Biochem. **72**, 19 (2003).
[39] J. Xing *et al.*, Proc. Natl. Acad. Sci. U.S.A. **103**, 1260 (2006).
[40] J. Tang and S. H. Lin, Phys. Rev. E **73**, 061108 (2006).
[41] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 545 (1997).
[42] S. S. Plotkin and J. N. Onuchic, Q. Rev. Biophys. **35**, 111 (2002).
[43] J. P. Ma, Curr. Protein Peptide Sci. **5**, 119 (2004).
[44] G. H. Li and Q. Cui, Biophys. J. **83**, 2457 (2002).
[45] R. Zwanzig, Proc. Natl. Acad. Sci. U.S.A. **85**, 2029 (1988).
[46] C. Hyeon and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **100**, 10249 (2003).

Allosteric effect: an alternative mechanism

Jianhua Xing[1, 2]

[1.] **Chemistry, Material and Life Science Directorate, University of California & Lawrence Livermore National Laboratory, Livermore, CA 94550**

[2.] Send correspondences to: xing3@llnl.gov

Abstract

Both the concerted MWC (Monod, Wyman, Changeux) and the sequential KNF (Koshland, Nemethy, Filmer) models of allosteric effects assume the thermodynamic properties of the allosteric enzymes are modified by substrate binding. Here we discussed an alternative mechanism that the dynamic properties of an allosteric enzyme being regulated by substrate binding. Theoretically it is possible that the reactivity of a protein is limited by some internal conformational change step (due to slow effective diffusion along rugged potential surfaces). Effactor binding may modify the ruggedness and thus the protein dynamics and reactivity. Compared to conventional models, the new mechanism has less requirements on the mechanical properties of an allosteric protein to propagate mechanical signals over long distances. Thus some signal transduction proteins may adopt the new mechanism or a combination with conventional mechanisms. The theory predicts that compared to the case with effactor binding, a positive allosteric enzyme alone under the new mechanism has: 1) larger temperature dependence; 2) larger effect of dynamic disorders; 3) smaller collective and individual dynamic fluctuations; 4) no requirement for a well-defined mechanical strain relaying network; 5) no requirement for large conformational change upon effactor binding. Opposite predictions of (1)-(3) are for negative allosteric proteins.

Introduction

A prominent property of proteins is that their catalytic activities can be regulated. Allosteric enzymes have two or more binding sites. Conformational changes due to ligand (effactor) binding or reaction on one site can propagate to another distant catalytic site and affect its reactivity. The discovery of allosteric regulations "in the 1950s, followed by a general description of allostery in the early 1960s, was revolutionary at the time"(1). There are two popular models proposed to explain the allosteric effects. The concerted MWC model by Monod, Wyman, and Changeux, assumes that an allosteric protein can exist in two (or more) conformations with different reactivity, and effactor binding modifies the thermal equilibrium distribution of the conformers(2). The sequential model described by Koshland, Nemethy, and Filmer is based on the induced-fit mechanism, and assumes that effactor binding results in (slight) structural change at another site and affects the substrate affinity(3). While different in details, both of the above models assume that the allosteric mechanism is through modification of the equilibrium conformation of the allosteric protein by effactor binding. For later discussions, we denote the mechanisms as "thermodynamic regulation". The above two mechanisms and generalizations have been used to explain observed allosteric effects. A notable extension is on explanation of the working mechanisms of protein motors. Protein motors convert chemical and electrochemical energies into mechanical work. Proper function of a protein motor requires elaborate regulation of chemical reaction steps and mechanical motion. We previously proposed that the above requirement is fulfilled with specific interactions forming potential bumps at various locations along

some mechanical (or conformational, see below) coordinates(4). Existence and removal of the potential bumps are regulated by substrate binding and dissociation. The mechanisms of thermodynamic regulation impose strong requirements on the mechanical properties of an allosteric protein. The distance between the two binding sites of an allosteric protein can be far. For example, the bacterial chemotaxis receptor has the two reaction region separated as far as 15 nm(5). The signal propagation requires a network of mechanical signal relaying residues with mechanical properties distinguishing them well from the surroundings to minimize thermal dissipation – Notice that distortion of a soft donut at one side has negligible effect on another side of the donut. Mechanical stresses due to effactor molecule binding irradiate from the binding site, propagate through the relaying network, and con-focus on the reaction region at the other side of the protein. It is challenging to transmit the mechanical stress faithfully against thermal dissipation over a long distance. A possible solution is the attraction shift model proposed by Yu and Koshland(6).

In this work, we will discuss a theoretically possible alternative mechanism of allosteric regulation, “dynamic regulation”. Instead of large change of conformation and thermodynamic properties such as substrate binding affinity, the protein activity can be modulated by modifying protein dynamic properties. This idea is inspired by experimental and theoretical studies on dynamic disorders. Dynamic disorders refer to the phenomena that the ‘rate constant’ of a process is actually a random function of time (7, 8). Since the pioneering work of Frauenfelder and coworkers on ligand binding to myoglobin(9), extensive experimental and theoretical studies have been performed in this subject (see for example ref.(8) for further references). Recently existence of dynamic

disorders has been demonstrated directly through single molecule enzymology measurements(10-14).

Theory and numerical results

The catalytic site of the protein can be described by a few numbers of reduced conformational degrees of freedom and the reaction coordinates. Here we use a minimal three-state model to represent the catalytic site chemical states: E (empty), R (reactant bound), P (product bound). An example used in this work is shown in Figure 1a.

The system dynamics can be well described by a set of over-damped generalized Langevin equations coupled to Markov chemical transitions (15),

$$0 = -\frac{dU_i}{dx} - \zeta_i \int_0^t d\tau M_i(t-\tau)\frac{dx(\tau)}{d\tau} + f_i(t), \quad (1)$$

where x represents the conformational coordinate[1], U_i is the potential of mean force at a given substrate binding state, ζ_i is the drag coefficient, M is the memory kernel, and f is the random fluctuation force. Chemical transitions accompany motions along the conformational coordinate with x-dependent transition rates. Effactor binding at a remote site can affect the dynamics at the catalytic site by modifying U_i, which is discussed by the conventional models of "thermodynamic regulation", the drag coefficient ζ_i, and/or the memory kernel M. The latter two are for "dynamic regulation". Our recent theoretical analysis showed that the observed slow protein conformational dynamics can be

[1] For simplicity, here we consider a one-dimensional case, which may represent, for example, the distance between two residues (see 16. Xing, J. & Kim, K. S. (2006) *Phys. Rev. E* **74,** 061911.).

explained by the rugged protein potential surfaces(16). For a potential surface with random ruggedness, Zwanzig showed that the effective diffusion constant is reduced by $D = D_0 \exp\left(-(\varepsilon / k_B T)^2\right)$, where D_0 is the bare diffusion constant, and ε is the potential roughness parameter(17). The reported values of ε is 2-6 k_BT(18, 19). With $D_0 = 10^{-6}$ cm^2/s, D can be reduced to 1 $Å^2/s$ with $\varepsilon \sim 4.8\ k_BT$. Thus internal diffusion can be a rate limiting step for enzymatic reactions and can in principle be regulated by allosteric effect (see Figure 1).

In this study we will focus on the case that the memory kernel is short (a δ-function, the so-called Langevin dynamics) and show that varying ζ_i alone can regulate protein activity. In this case, the dynamics at steady state can be equally described by a set of coupled Fokker-Planck equations,

$$-\frac{D_i}{k_B T}\cdot\frac{\partial}{\partial x}\left(-\frac{\partial U_i(x)}{\partial x}\rho_i\right)+D_i\frac{\partial^2\rho_i}{\partial x^2}+\sum_{j\neq i}\left(K_{ij}(x)\rho_j-K_{ji}(x)\rho_i\right)=0\ ,(2)$$

Where k_B is the Boltamann's constant, T the temperature, $D_i = k_BT/\zeta_i$ is the diffusion constant, K_{ij} is the transition matrix element, and ρ_i is the probability density to find the system at position x and state i. In our numerical calculations, the potentials are chosen as harmonic potentials, $V_i = \frac{1}{2}\kappa_i\left(x - x_{0i}\right)^2 + V_{0i}$, $i = (E, R, P)$. To model transitions between different states, we also model the transition state potentials by harmonic potentials, $V_{ij}^{‡} = \frac{1}{2}\left((x - x_{ij}^c)/L_{ij}\right)^2 + V_{ij}$. The transition rate from state j to i is given by $k_{ij}(x) = k_{ij}^0 \exp\left[\left(V_j(x) - V_{ij}^{‡}(x)\right)/k_BT\right]$. In general the optimal transition location is different for different reactions, and some conformational motion is needed during an enzymatic cycle. Model parameters are given in Transition parameters

	R←E	P←E	E←R	P←R	E←P	R←P
k^0	2e2	2e-3	2e2	1.6e3	2e3	1.6e3
$V_0^{\dagger}$	3	3	3	6	3	6
L	0.3	0.3	0.3	0.3	0.3	0.3

Table 1.

The Fokker-Planck equations were solved with the algorithm developed by Wang *et al*(20). Figure 2a shows the calculated enzyme turnover rate as a function of the internal diffusion constant. While the diffusion constant is not a rate-limiting parameter at high values (as compared to the chemical transition rates), at smaller values of D the turnover rate depends on the diffusion constant linearly which signatures existence of diffusion-limited steps. Figure 2b shows the temperature dependence of the turnover rate. With high values of D, the exponential $1/T$ dependence mainly comes from the Arrhenius dependence of the transition rates. However with small values of D, the turnover rate shows strong non-exponential dependence, since the effective diffusion constant D has a Gaussian dependence on $1/T$. Therefore the theory predicts strong non-exponential temperature dependence of enzyme activity in the absence of the effactor if it is regulated by the dynamic mechanism. Figure 2c also shows the waiting time distribution between two consecutive turnover cycles. The results are calculated with the formula derived by Gopich and Szabo(21). A system with low D values shows non-exponential distribution due to dynamic disorders. At high D values, effects of the dynamic disorders diminish and the distribution is exponential. Therefore, we predict that an enzyme functioning under the dynamic regulation mechanism shows larger dynamic disorders effect. This can be directly tested by measuring consecutive single enzyme turnover time distributions with and without the effactor, an extension of the work done by the Xie group(13).

Discussions

In short, here we propose that protein dynamics can be rate-limited by some internal diffusion steps, and effactor binding can change the situation by modifying the effective internal diffusion constant. Compared to conventional thermodynamic regulation mechanism, the dynamic regulation mechanism has the following advantages.

First, it is an effective way to regulate the reactivity at a distant region. For an Arrhenius process, to increase the reactivity by 10^{10}, the activation barrier needs to be lowered by 23 k_BT. On the other hand, for an internal diffusion limited process, the reaction rate is linearly dependent on the effective diffusion constant. The lower bound of the roughness parameter adjustment is 4.8 k_BT to increase the reactivity by 10^{10}.

Second, it has less requirement on the mechanical properties of the protein. Figure 3a schematically shows the free energy profiles of a protein with and without the effactor under the conventional allosteric mechanism. The free energy difference between the two curves, $\Delta U(x)$, is provided by the effactor binding energy as a function of the conformational coordinate. Effective coupling of the two sites requires faithful transmission of the mechanical strain due to ligand binding from one site to another one. In terms of solid mechanics, the mechanical stress lines should propagate from one site and concentrate on another site. A set of mechanical stress relaying network is expected to perform the task. These network residues should have mechanical properties distinctive from other residues to minimize lost to the surroundings (see Figure 3b). In other words, coupling between these relaying residues and others should be minimized. By comparison, under a dynamics regulation mechanism, the effect can be highly

delocalized. Effactor binding may affect the other site by fine regulating local structures far away from that site. By modifying the effective diffusion constant, these local modifications may affect the dynamics of collective motions including residues within the catalytic site and those far from it. These collective motions affects enzyme reactivity by projecting onto the conformational coordinate x, which is directly coupled to the chemical reaction coordinate (see Figure 1a and ref. (16)). The effect manifests itself through lager root-mean-square deviation as observed in NMR, x-ray crystography, and in molecular dynamics simulations. We suggest the dynamics of the collective motions to be examined as well.

Recently relations between protein dynamic properties and allosteric effects have been extensively studied by both experimental and simulation techniques(5, 22-24). It is proposed that the entropic changes associated with ligand binding contribute to the allosteric effect. This mechanism still falls into the category of the "thermodynamic regulation" discussed in this paper, since its effect is to modify the free energy function U(x). The mechanism discussed in this work is different. Some of the experimental evidence in supporting of the dynamic entropic effect mechanism, may also be explained by the dynamic mechanism discussed in this work. For example, both suggest a distribution of protein conformations, and possibility of lacking well-defined mechanical strain relaying network. The allosteric mechanism of a given protein likely has contribution from both thermodynamic regulation (the conventional conformational change mechanism and the newly proposed entropic effect), and dynamic regulation proposed in this work (see also Equation (1)).

Acknowlegements

JX is supported by a Lawrence Livermore National Laboratory Directed Research and Development grant, and by a Chemistry, Material, and Life Sciences Directorate fellowship. This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

References

1. **Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002)** ***Molecular Biology of the Cell*** **(Garland, New York).**
2. **Monod, J., Wyman, J. & Changeux, J. P. (1965)** ***J. Mol. Biol.*** **12, 88-&.**
3. **Koshland, D. E., Nemethy, G. & Filmer, D. (1966)** ***Biochemistry*** **5, 365-&.**
4. **Xing, J., Liao, J.-C. & Oster, G. (2005)** ***Proc. Natl. Acad. Sci. USA*** **102, 16539-16546.**
5. **Kim, S. H., Wang, W. R. & Kim, K. K. (2002)** ***Proc. Natl. Acad. Sci. U.S.A.*** **99, 11611-11615.**
6. **Yu, E. W. & Koshland, D. E. (2001)** ***Proc. Natl. Acad. Sci. U.S.A.*** **98, 9517-9520.**
7. **Frauenfelder, H., Wolynes, P. G. & Austin, R. H. (1999)** ***Rev. Mod. Phys.*** **71, S419-S430.**
8. **Zwanzig, R. (1990)** ***Acc. Che. Res.*** **23, 148-152.**
9. **Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. (1975)** ***Biochemistry*** **14, 5355-5373.**
10. **Xie, X. S. & Lu, H. P. (1999)** ***J. Biol. Chem.*** **274, 15967-15970.**
11. **Kou, S. C., Cherayil, B. J., Min, W., English, B. P. & Xie, X. S. (2005)** ***J. Phys. Chem. B*** **109, 19068-19081.**
12. **Min, W., English, B. P., Luo, G. B., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2005)** ***Acc. Chem. Res.*** **38, 923-931.**
13. **English, B. P., Min, W., van Oijen, A. M., Lee, K. T., Luo, G. B., Sun, H. Y., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2006)** ***Nat. Chem. Biol.*** **2, 87-94.**
14. **Min, W., Luo, G. B., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2005)** ***Phys. Rev. Lett.*** **94, 198302-.**
15. **Zwanzig, R. (2001)** ***Nonequilibrium Statistical Mechanics*** **(Oxford University Press, Oxford).**
16. **Xing, J. & Kim, K. S. (2006)** ***Phys. Rev. E*** **74, 061911.**
17. **Zwanzig, R. (1988)** ***Proc. Natl. Acad. Sci. U.S.A.*** **85, 2029-2030.**
18. **Nevo, R., Brumfeld, V., Kapon, R., Hinterdorfer, P. & Reich, Z. (2005)** ***EMBO Rep.*** **6, 482-486.**
19. **Thirumalai, D. & Woodson, S. A. (1996)** ***Accounts of Chemical Research*** **29, 433-439.**

20. **Wang, H., Peskin, C. & Elston, T. (2003)** ***J. Theo. Biol.*** **221, 491-511.**

21. **Gopich, I. V. & Szabo, A. (2006)** ***J. Chem. Phys.*** **124, -.**

22. **Kern, D. & Zuiderweg, E. R. P. (2003)** ***Curr. Opin. Struc. Biol.*** **13, 748-757.**

23. **Formaneck, M. S., Ma, L. & Cui, Q. (2006)** ***Proteins-Structure Function and Bioinformatics*** **63, 846-867.**

24. **Fuentes, E. J., Gilmore, S. A., Mauldin, R. V. & Lee, A. L. (2006)** ***J. Mol. Biol.*** **364, 337-351.**

Potential parameters

	κ	x_0	V_0
E	1	1	0
R	0.5	-0.5	-1
P	0.4	0	-3

Transition parameters

	R←E	P←E	E←R	P←R	E←P	R←P
k^0	2e2	2e-3	2e2	1.6e3	2e3	1.6e3
$V_0^{\dagger}$	3	3	3	6	3	6
L	0.3	0.3	0.3	0.3	0.3	0.3

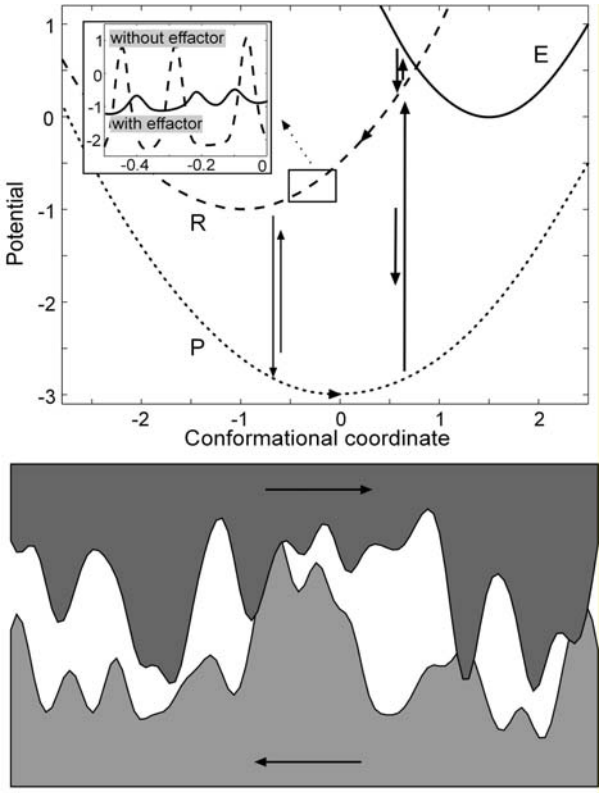
Table 1 Model parameters. All are in reduced units: for energy $k_BT = 1$ except the temperature dependence results in Figure 2. For simplicity, the diffusion constants for the three states take the same value.

Figure captions

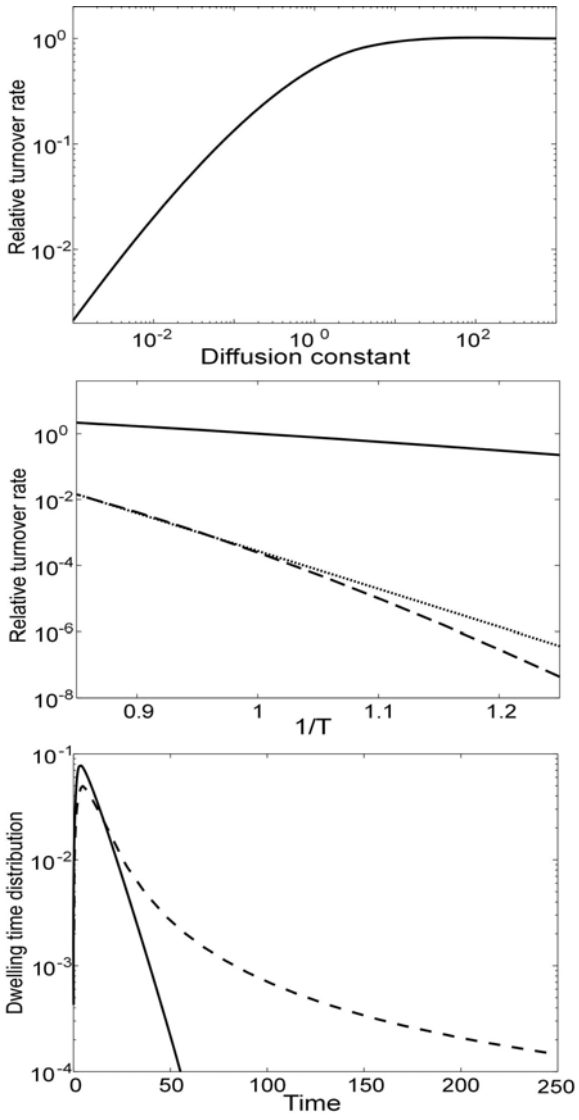
Figure 1 (a)Three free energy curves represent three distinct catalytic site binding states (E, R, P) along some conformational coordinate x. Chemical transitions between the three states are centered at some x values. The inlet illustrates that the smooth potentials are actually coarse-grained over rugged potential surfaces. Effactor binding at a distant site can modify the roughness of the potentials. (b) Physically one source of the potential roughness comes from the tightly packed irregular surfaces between two moving parts inside a protein.

Figure 2 (a) Enzyme turnover rate as a function of the effective internal diffusion constant D. (b) Temperature dependence of the enzyme turnover rate. $D = D_0 \exp[-(\varepsilon/k_BT)^2]$, where ε is the roughness parameter, and $D_0 = 10^3$. Solid line: ε = 0. dashed line: $\varepsilon = 4\ k_BT$. The temperature dependence of D_0 is neglected in this calculation. (c) Turnover waiting time distribution with D = 1 (solid line) and D = 0.001 (dashed line).

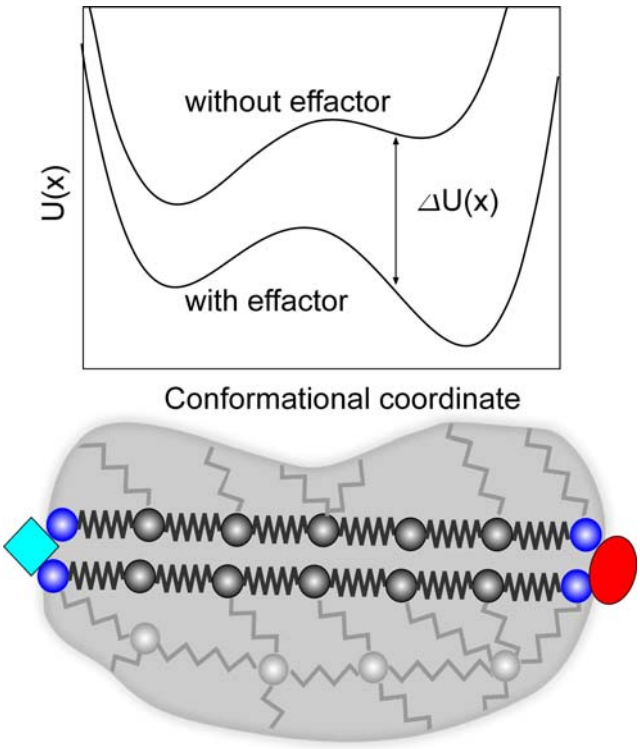
Figure 3 schematic illustration of the conventional 'thermodynamic regulation. (a) The free energy profile of the protein as a function of some conformational coordinate with and without the effactor binding. Different models (MWC, KNF, etc) differ in some details of the shapes (e.g, single or multiple minima). (b) Effective coupling of the two binding sites require a set of residue network to transmit mechanical stress between the two sites despite of thermal dissipation. These residues are expected to have solid-like properties, and are distinctive from the surrounding residues.



Xing Figure 1



Xing Figure 2



Xing Figure 3

Protein switch: detailed mechanisms and effects of dynamic disorders

Jianhua Xing[1, 4], Jing Chen[2], Wei Min[3], Sunney X. Xie[3]

[1.] Chemistry, Materials, and Life Sciences Directorate, University of California & Lawrence Livermore National Laboratory, Livermore, CA 94550

[2.] Department of Molecular Cell Biology, University of California, Berkeley, CA 94720

[3.] Department of Chemistry and Chemical Biology, Harvard University,

[4.] Send correspondences to: xing3@llnl.gov

Total 14 pages, 15782 characters, 4 figures, 1 table

Abstract

Many existing mathematical models describing protein interaction networks involve phenomenological (e.g., Michaelis-Menten type) rate equations. In this work we studied the classical example of protein switch through covalent modification first studied by Goldbeter and Koshland. We pointed out that the switching behavior cannot be derived from their model. The discrepancy is due to misuse of the Michaelis-Menten approximation in the original work. We examined some rescuing schemes. We also showed that a covalent modification protein switch is robust to the existence of dynamic disorders at bulk concentration. However, if the dynamic disorder is quasi-static, large fluctuations of the switch response behavior may be observed at low concentrations relevant to many biological functions. This source of network dynamics fluctuation is different from the types of fluctuations usually discussed in the literature. Fluctuations of the switch behavior can be reduced by increasing the protein conformer inter-conversion rate, or by correlating the enzymatic reaction rates in the network.

Introduction

A biological system usually functions by regulating protein activities through protein interaction networks (PINs), which demonstrate rich dynamic behaviors. Substantial mathematical modeling has been carried out to study these network regulations. Some typical motifs of PINs were summarized in a recent review by Tyson *et. al*(1). Many of the existing mathematical models describe the dynamics of the PINs with phenomenological rate equations, like the Michaelis-Menten equations. These phenomenological equations imply asymptotic assumptions about the reaction rates and the concentration of reactants. However, these assumptions are not always valid. When these assumptions are lifted, the unwrapped PIN models defined with simple chemical kinetics may display significantly different dynamic behaviors from what their phenomenological counterparts do (Sabouri-Ghomi *et. al*, preprint). Therefore, extra reaction steps are necessary to reproduce the desired dynamic behaviors that the phenomenological models yield, such as oscillations and switches. In this work we will focus on a classical example of protein switch first studied by Goldbeter and Koshland(2).

We also studied the effect of dynamic disorders on the switch. Dynamic disorders refer to the phenomena that the 'rate constant' of a reaction appears as a random function of time (3, 4). This phenomena has attracted extensive experimental and theoretical studies ever since the pioneering work by Frauenfelder and coworkers on ligand binding to myoglobin(5). Recently, the existence of dynamic disorders has been directly confirmed by single-molecule enzymological measurements(6-8). We refer the readers to some review articles and references therein for more information(9-12). Physically, the existence of dynamic disorders comes from coupling between the reaction coordinate and other slowly fluctuating protein conformational coordinates. Due to ruggedness of protein potential landscapes, the time scales of protein conformational fluctuations can be comparable or even slower than the chemical

reaction time scale(9, 10). Chemical rate theories are usually based on the assumption of time scale separation between a slow reaction coordinate dynamics and fast conformational coordinate dynamics. This assumption is generally not valid in protein dynamics(13). Consequently, a rate 'constant' for an enzymatic reaction is actually a function of protein conformations.

Detailed mechanisms of the sigmoidal switch

We have chosen to study the sigmoidal switch, one of the most basic motifs of PINs. In their classical work, Goldbeter and Koshland considered a covalent modification system, typically composed of phosphorylated-dephosphorylated couples: $E \underset{\uparrow A}{\overset{\downarrow S}{\rightleftharpoons}} EP$. Goldbeter and Koshland assumed that both the phosphorylation and dephosphorylation reactions obey Michaelis-Menten dynamics,

$$\begin{aligned} \frac{d}{dt}[E] &= -\frac{k_1[S]\left([E_t]-[EP]\right)}{K_{m1}+[E_t]-[EP]} + \frac{k_2[EP]}{K_{m2}+[EP]}, \\ \frac{d}{dt}[EP] &= \frac{k_1[S]\left([E_t]-[EP]\right)}{K_{m1}+[E_t]-[EP]} - \frac{k_2[EP]}{K_{m2}+[EP]}, \end{aligned} \qquad (1)$$

where K_{m1} and K_{m2} are the two Michaelis-Menten constants. In the above expressions, [X] represents the concentration of the reactant X. The steady state concentration of the phosphorylated form [EP] as a function of the signal concentration [S] is given by the so-called Goldbeter-Koshland function, which shows a switch-like behavior called "zero-order ultrasensitivity" in its responses to the varying signal level(2), characterized by a sigmoidal signal-response curve (Figure 2a). Such sigmoidal switch also plays a role of building blocks in other more complex PINs. For example, it allows phase offset between X and R in the negative feedback oscillator example discussed by Tyson *et. al*.(1). Therefore, it is a good starting point to study PINs.

In this work, we conducted both numerical and analytical computation on the dynamics of the covalent modification system with explicit consideration of the intermediate complexes neglected in the Michaelis-Menten treatment. The governing equations are (see case a of Figure 1),

$$\frac{d}{dt}\begin{pmatrix}[E]\\[ES]\\[EP]\\[AEP]\end{pmatrix}=\begin{pmatrix}-k_{1f}S & k_{1r} & 0 & k_2\\ k_{1f}S & -k_{1r}-k_1 & 0 & 0\\ 0 & k_1 & -k_{2f} & k_{2r}\\ 0 & 0 & k_{2f} & -k_{2r}-k_2\end{pmatrix}\begin{pmatrix}[E]\\[ES]\\[EP]\\[AEP]\end{pmatrix} \tag{2}$$

with

$$[E]+[ES]+[EP]+[AEP]=[E_t],\ [S]+[ES]=[S_t] \tag{3}$$

We assume that A, the dephosphatase, is in such great excess that its concentration, [A], remains approximately constant throughout time, and is absorbed in k_{2f} for mathematical simplicity. Similar assumption is made throughout the paper. Relaxing this assumption does not qualitatively change the conclusion of this paper. Unlike the behavior of its phenomenological counterpart, throughout the parameter space this model shows no sigmoidal response either to the changes in [S], the free signal concentration, or to the changes in [S_t], the total signal concentration. (see Figure 2b). This observation is confirmed by the analytic result that the second derivatives of [EP] with respect to [S] and [S_t] are both identically negative, i.e.

$$\frac{d^2}{dS^2}[EP]<0,\frac{d^2}{dS_t^{\,2}}[EP]<0,\ \text{for [S], [S}_t\text{]}>0 \tag{4}$$

We considered a scheme (case a2) with an additional step: E and S firstly form a weakly bound form, ES, then convert to the tightly bound form, E*S, which proceeds to the phosphorylation step. The desired sigmoidal response curve does not emerge either. Details of the analysis are given in the Supporting Text.

The above analysis suggests that nonlinear terms of [S] are necessary to generate the sigmoidal behavior. We examined one such scheme (case b, Figure 1), inspired by the work of Sabouri-Ghomi *et. al* (preprint). In this case, we have modified the model such that binding an additional S molecule to ES, the intermediate compound, facilitates the phosphorylation reaction,

$$\begin{pmatrix} -k_{1f}[S] & k_{1r} & 0 & k_2 & 0 \\ k_{1f}[S] & -k_{1r}-k_1-k_{1f}^{'}[S] & 0 & 0 & k_{1r}^{'} \\ 0 & k_1 & -k_{2f} & k_{2r} & k_1^{'} \\ 0 & 0 & k_{2f} & -k_2-k_{2r} & 0 \\ 0 & k_{1f}^{'}[S] & 0 & 0 & -k_{1r}^{'}-k_1^{'} \end{pmatrix} \begin{pmatrix} [E] \\ [ES] \\ [EP] \\ [AEP] \\ [SES] \end{pmatrix} = 0 \qquad (5)$$

with

$$\begin{aligned} &[E]+[ES]+[EP]+[AEP]+[SES]=[E_t], \\ &[S]+[ES]+2[SES]=[S_t] \end{aligned} \qquad (6)$$

The additional mechanism incurs a nonlinear term of [S] (see Equation 6) and yields the desired sigmoidal response (Figure 2). Figure 2b gives concentrations of various enzyme forms as a function of $[S_t]$ in case b. [ES] increases steadily with initial increase of [S], then decreases suddenly. In other words, for low concentrations of $[S_t]$, a large proportion of the signal molecule is concealed in the bound form, ES. Only when $[S_t]$ reaches a critical value, a large portion of the protein is transformed in the form of [EP] via the intermediate [SES].

For case b, we also examined the situation that the free signal concentration [S] instead of $[S_t]$ is controlled. In this case our mathematical analysis (see Supporting Text for details) shows that [EP] as a function of [S] can only yield mild sigmoidal response. Numerical searching of the parameter space confirmed this conclusion.

Effects of dynamic disorders

In this part, we will describe the effect of dynamic disorders on the sigmoidal switch, in particular, on the network presented in case b. We adopted a multistate model studied by Kou *et. al*(14). The model assumes *N* conformations of the enzyme, E_1, …, E_N. Corresponding *N* conformations exist for all the other states of the enzyme, for example, ES_1,…, ES_N for ES. Only matching conformations of the reactant and the product are admissible in chemical transitions, e.g. $EP_2 \rightarrow AEP_2$, but not $EP_2 \rightarrow AEP_3$. Part or all of the chemical transition rates vary with different conformations, so the rates of individual enzymes display temporal fluctuation as the enzyme randomly converts to different conformations. Dynamic disorders thus appear. Here we assumed that the rates of the disordered chemical transitions obey a gamma distribution over all conformations with certain mean and variance (with the mathematical expression given in the caption of Figure 3). Interconversions occur between each pair of conformers of the same enzyme state, not restricted to the neighboring numbered ones, $X_i \leftrightarrow X_{i+1}$. For simplicity, a uniform rate was used for all the conformer interconversions.

In Figure 3, we compared the ensemble averaged signal-response curve, [EP] vs $[S_t]$, with the signal-response curve of each conformation, $[EP]_i$ vs $[S_t]_i$. Here the ensemble concentration of the signal and the response are defined as

$$[EP] = \sum_{i=1}^{N} [EP]_i, [S_t] = [S] + \sum_{i=1}^{N} \left([ES]_i + 2[ES \bullet ES]_i \right) \qquad (7)$$

And the signal concentration for a given conformation assumes a scaled form $[S_t]_i = [S] + N\left([ES]_i + 2[ES \bullet ES]_i \right)$. In the limit of quasi-static protein conformational changes, the response curve of a given conformation *i* is virtually the response relation with existence of only conformer *i*.

Figure 3a shows that the ensemble response curve is not significantly affected by the conformer conversion rate. Similar results are obtained by imposing dynamic disorders on other enzymatic reaction steps. Therefore, the effect of dynamic disorder on the system response is barely noticeable if the response curve is measured at bulk. However, large fluctuations in the system response becomes significant if the protein concentration [E_t] is low (e.g., close to single molecule level). Figure 3b and Figure 4a show that variance in k_1' (or k_2) results in variance of the sigmoidal transition location up to 20%. Variance in k_1 causes much less fluctuations in the response because k_1 is not associated with the main reaction pathway in the network. The effect of dynamic disorders is dramatically reduced with increasing conformer inter-conversion rate (Figure 3c and Figure 4), or with perfect correlations between the disordered reaction rates (associated with k_1, k_2, and k_1', black dash line in Figure 4). The latter observation is because [EP] depends mostly on the ratio of three rate constants k_1, k_2, and k_1' (see the analytic result of case b in Supporting Text).

As discussed in Introduction, enzymatic reaction rates are in general multi-dimensional functions of protein conformations. By using correlated rates as above, we implicitly assume that these rates share similar function forms of protein conformational coordinates (Supporting figure

Figure 5a). Generally speaking, different reactions within a reaction network may have different dependence on protein conformational coordinates (Supporting figure

Figure 5b), and the dynamic disorder will not be offset. The red dashed line in Figure 4 shows that large variance exists even when k_1' just has a broader distribution than k_1 and k_2.

Discussions

This work has two focuses. First we discussed the detailed mechanism of the switch module. We showed mathematically that the covalently-modification system studied by Goldbeter and Koshland cannot produce the switch-like "zero-order ultrasensitivity" behavior with its original form. Rather, some extra steps are needed. We examined one such scheme by allowing enzymatic reaction rates affected by formation of complexes like [SES]. Given that the switch module is a basic PIN motif, further experimental studies on testing the theoretical results here should be necessary. We want to point out that the theoretical analysis presented in this work is based on the assumption that the law of mass action is valid. Violation of this assumption may lead to new mechanisms not discussed here.

Then we discussed the effect of dynamic disorders on the dynamic behaviors of the switch module. The ensemble averaged behavior is insensitive to the existence of dynamic disorders. However, fluctuations of the module response may be observed with a small number of protein molecules. There are several ways of reducing the effect of dynamic disorders (and thus increase robustness of the network): increasing the conformer inter-conversion rate, or correlating conformation fluctuation within the PIN. These mechanisms may have been adopted along evolution. However, dynamic disorders may even play a positive role. Tănase-Nicola et. al showed that noise correlations within a biochemical network can reduce the overall noise of the network(15). Existence of dynamic disorder is one noise source which can have broad time scales, and could interplay with other sources of noise.

Computational details

For the sigmoidal switch model, two optimization methods were used to search the parameter space: the downhill simplex optimization method, and simulated annealing. With an initial set of rate constant values, both the two algorithms search the parameter space to minimize the function

$$\int_0^1 \left|[RP(S)] - T(S)\right| dS \approx \frac{1}{N}\sum_{i=1}^{N}\left|[RP(S_i)]T(S_i)\right| \quad (8)$$

where $[RP(S_i)]$ is the calculated RP concentration as a function of the signal strength, and T(S) is the desired signal-response curve. In all the calculations T(S) assumes the 'Goldbeter-Koshland' function form,

$$G(u,v,J,K) = \frac{2uK}{v-u+vJ+uK+\sqrt{\left(v-u+vJ+uK\right)^2 - 4\left(v-u\right)uK}} \quad (9)$$

with $u = 2S, v = 1, J = K = 0.05$.

With existence of dynamic disorders, the enzymatic reaction rates assume a gamma distribution

$$p(k) = [1/(b^a\Gamma(a))]k^{a-1}\exp(-k/b)$$

In our model calculations, a number of N conformers are used to model a given protein binding state. First the k coordinate is divided into N regions with the end points k_i^e chosen so that

$$\int_{k_{i-1}^e}^{k_i^e} p(k)dk = 1/N, i = 1,\ldots,N \quad (10)$$

and $k_0^e = -\infty, k_N^e = \infty$. Then the N discretized rate constants are chosen as,

$$k_i = \frac{\int_{k_{i-1}^e}^{k_i^e} dk\ p(k)k}{\int_{k_{i-1}^e}^{k_i^e} dk\ p(k)} = N\int_{k_{i-1}^e}^{k_i^e} dk\ p(k)k, i = 1,\ldots,N \quad (11)$$

Acknowledgements

JX is supported by a Lawrence Livermore National Laboratory Directed Research and Development grant, and by a Chemistry, Material, and Life Sciences Directorate fellowship. JC is supported by … . MW and SX are supported by … . This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

References

1. Tyson, J. J., Chen, K. C. & Novak, B. (2003) *Curr. Opin. Cell Biol.* **15,** 221-231.
2. Goldbeter, A. & Koshland, D. E. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78,** 6840-6844.
3. Frauenfelder, H., Wolynes, P. G. & Austin, R. H. (1999) *Rev. Mod. Phys.* **71,** S419-S430.
4. Zwanzig, R. (1990) *Acc. Che. Res.* **23,** 148-152.
5. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. (1975) *Biochemistry* **14,** 5355-5373.
6. Xie, X. S. & Lu, H. P. (1999) *J. Biol. Chem.* **274,** 15967-15970.
7. English, B. P., Min, W., van Oijen, A. M., Lee, K. T., Luo, G. B., Sun, H. Y., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2006) *Nat. Chem. Biol.* **2,** 87-94.
8. Min, W., Luo, G. B., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2005) *Phys. Rev. Lett.* **94,** 198302-.
9. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. (1991) *Science* **254,** 1598-1603.
10. Zwanzig, R. (1992) *J. Chem. Phys.* **97,** 3587-3589.
11. Min, W., English, B. P., Luo, G. B., Cherayil, B. J., Kou, S. C. & Xie, X. S. (2005) *Acc. Chem. Res.* **38,** 923-931.
12. Karplus, M. (2000) *J. Phys. Chem. B* **104,** 11-27.
13. Xing, J. & Kim, K. S. (2006) *Phys. Rev. E* **74,** 061911.
14. Kou, S. C., Cherayil, B. J., Min, W., English, B. P. & Xie, X. S. (2005) *J. Phys. Chem. B* **109,** 19068-19081.
15. Tanase-Nicola, S., Warren, P. B. & ten Wolde, P. R. (2006) *Phys. Rev. Lett.* **97,** 068102.

Table 1 Model parameters.

	k_1	k_{1f}	k_{1r}	$k_{1f}^{'}$	$k_{1r}^{'}$	$k_1^{'}$	k_2	k_{2f}	k_{2r}	E_t
Case a	5	0.06	20				1	0.06	20	1
Case b	0.006	400	12	40	670	10	1	4	79	1

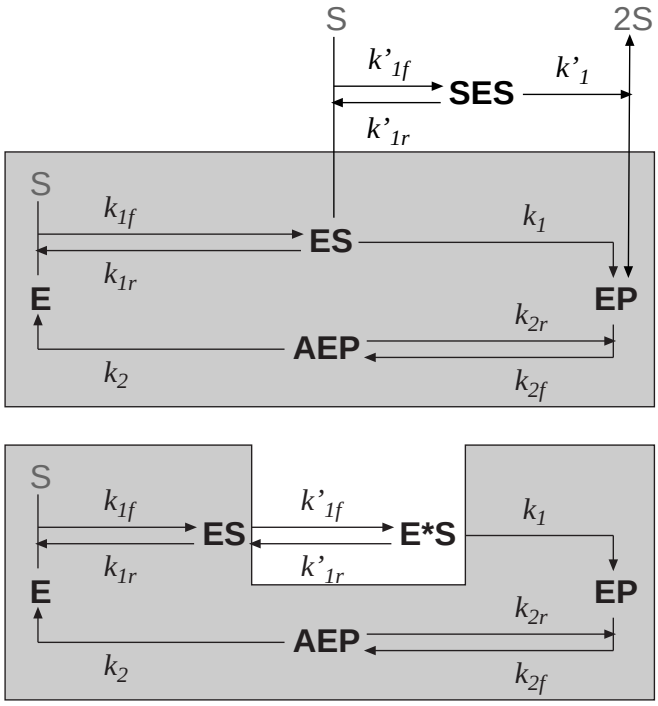
Figures

Figure 1 Several possible detailed unwrapped schemes of the sigmoidal switch studied by Goldbeter and Koshland. Case a (upper panel within the shaded box): the switch is composed of an enzyme in phosphorylated and dephosphorylated forms, with the transformation regulated by another signal molecule. Case a2 (lower panel): E and S form a loosely bound complex, then a tightly bond one preceding the phosphorylation step. Case b (upper panel with extra pathway shown in blue): ES can bind another S molecule, which accelerate the phosphorylation reaction.

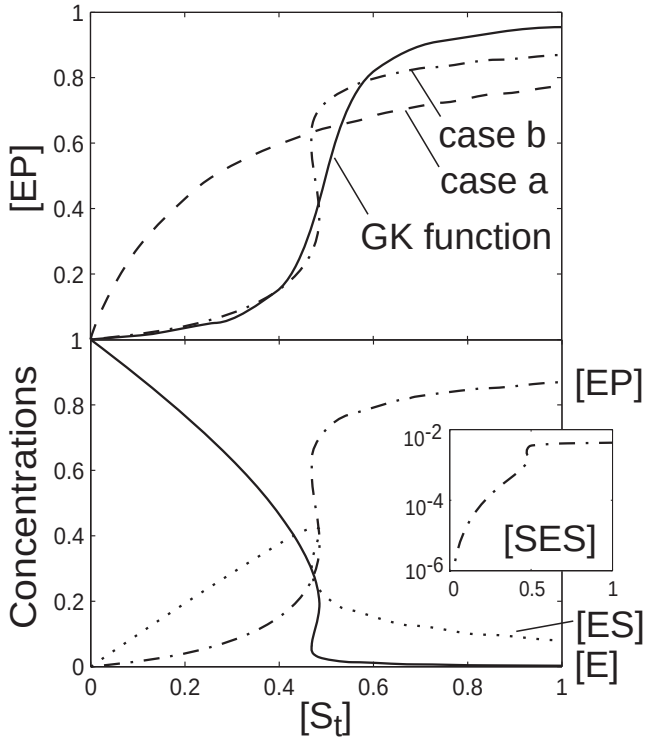
Figure 2 (a) Exampling signal-response curves obtained with the schemes discussed in Figure 1. The G-K function is the Goldbeter and Koshland result obtained by assuming both the phosphorylation and dephosphorylation steps are governed by Michaelis-Menten kinetics. Parameters of other cases are obtained by fitting the G-K function (see Supporting Text for details), and are given in Table 1. (b) Concentrations of various enzyme forms as a function of the total signal concentration in case b.

Figure 3 Response curves of case b in the presence of dynamic disorders of k_1'. For the results shown, k_1' assumes a gamma distribution $p(k) = [1/(b^a \Gamma(a))] k^{a-1} \exp(-k/b)$ with Γ(a) being the gamma function, and a = 4, b = 10/a (here 10 is the rate constant value used in the absence of dynamic disorder). The number of conformers is 10, with the discretization method discussed in ST. (a) The ensemble averaged response curves with the conformer interconversion rate $k_{ij} = 10^{-5}$ (solid line) and $k_{ij} = 10^{-1}$ (dashed line), respectively. (b) Contributions from individual conformers with $k_{ij} = 10^{-5}$ (dotted lines). The ensemble averaged result is also shown (solid line). (c) Same as b except $k_{ij} = 10^{-1}$. The quantities σ_S and σ_{EP} are the relative variances shown in Figure 4.

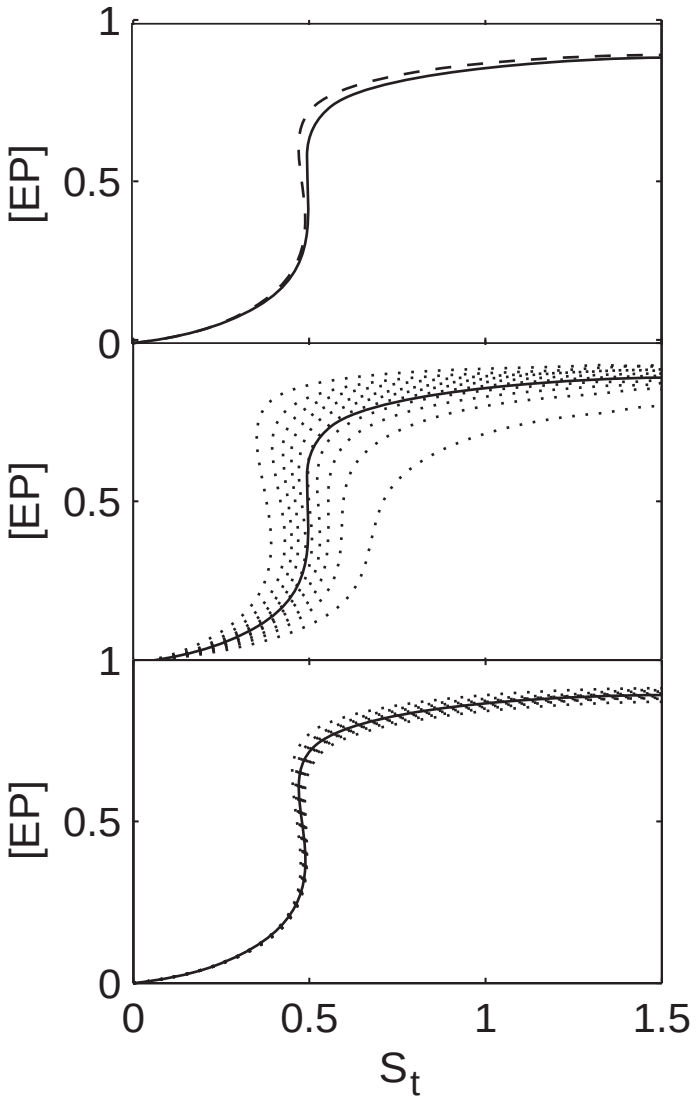
Figure 4 Upper: relative variance of $[S_t]_i$ at $[EP]_i = [EP(S\rightarrow\infty)]/2$ as a function of the conformer inter-conversion rate. Bottom: relative variance of $[EP(S\rightarrow\infty)]_i$ as a function of the conformer inter-conversion rate. a) k_2 has a gamma distribution with $a_2 = 4$, $b_2 = 1/a_2$, (thus the average of k_2 is 1, the value used in the case without dynamic disorder, see Table 1); b) k_1' has a gamma distribution with $a_1' = 4$, $b_1' = 10/a_1'$; c) k_1 has a gamma distribution with $a_1 = 4$, $b_1 = 0.008/a_1$; d) all the enzymatic reaction rates, k_1, k_1', and k_2, have gamma distribution with (a_1, b_1; a_1', b_1'; a_2, b_2); e) k_1, and k_2, have gamma distribution with a_1, b_1; a_2, b_2), but k_1' has a gamma distribution with (2, 10/2).



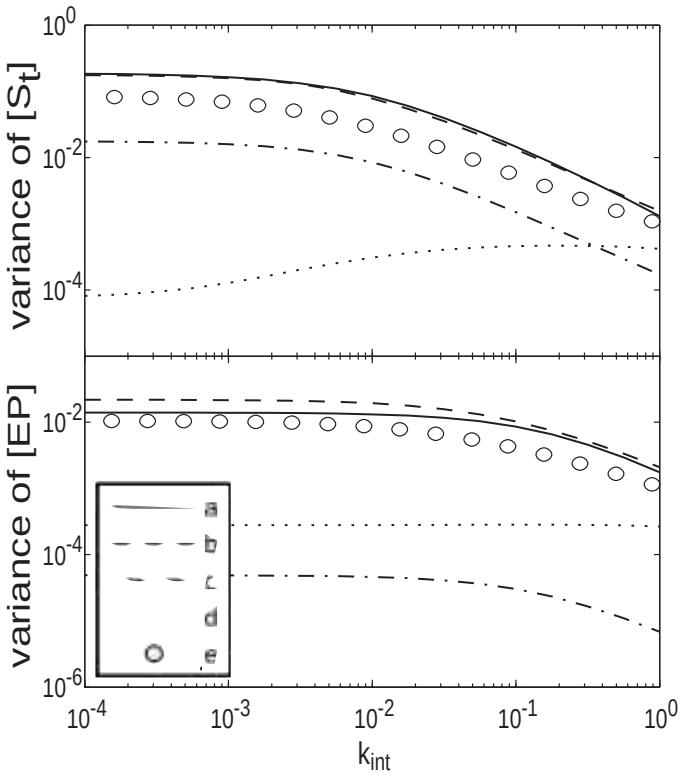
Xing et al. Figure 1



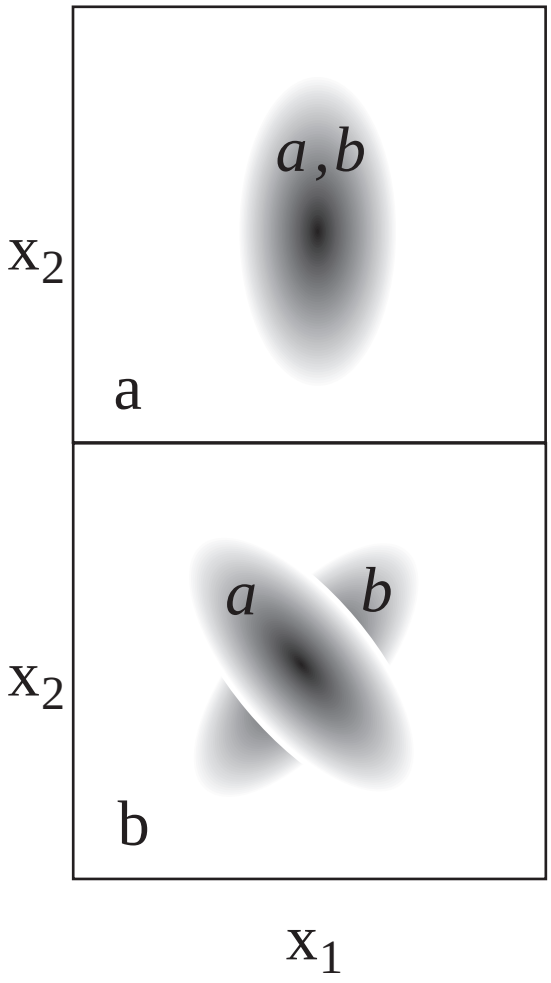
Xing et al. Figure 2



Xing et al. Figure 3



Xing et al. Figure 4



Xing et al. Figure 5

Supporting figure

Figure 5 Schematic illustration that rate changes due to conformational fluctuations may not correlate for two substrate binding states. Here x_1 and x_2 represent two (of the many) protein conformational coordinates. An enzymatic reaction rate *k* is a function of (x_1, x_2), which is represented by the shaded region. A darker region means higher reaction rate. a: If for two binding states, the two functions overlap well, the two rates are correlated, and a higher k_α corresponds to a higher k_β. Then the effect of dynamic disorders is greatly reduced even for the quasi-static case. b: In general, the two functions do not overlap well, and large effect of dynamic disorders may be observed for the covalent modification switch.

Supporting Text

A. Impossibility of sigmoidal behaviors in case a and a2

Case a

We examine the steady-state solution of Equation (2) and (3) in the main text. First other quantities can be expressed by [E],

$$[ES] = \frac{[S_t][E]}{K_{m1} + [E]}$$
$$[AEP] = \frac{k_1}{k_2}[ES] = \frac{k_1}{k_2}\frac{[S_t][E]}{\left(K_{m1} + [E]\right)} \quad (1)$$
$$[EP] = K_{m2}[AEP] = \frac{k_1}{k_2}\frac{K_{m2}[S_t][E]}{\left(K_{m1} + [E]\right)}$$

Then the physically meaningful solution of [E] is

$$[E] = \frac{-B + \sqrt{B^2 + 4[E_t]K_{m1}}}{2}$$
$$[EP] = \frac{k_2 K_{m2}\left(a[S_t] + K_{m1} + [E_t] - \sqrt{a^2[S_t]^2 + 2ab[S_t] + \left([E_t] + K_{m1}\right)^2}\right)}{2a} \quad (2)$$

Where,

$$K_{m1} = \frac{k_{1r} + k_1}{k_{1f}}, K_{m2} = \frac{k_{2r} + k_2}{k_{2f}},$$
$$B = \left(1 + k_2 + k_2 K_{m2}\right)[S_t] + K_{m1} - [E_t] = a[S_t] + b$$

Then

$$\frac{2a}{k_2 K_{m2}} \frac{d}{d[S_t]}[EP] = a - \frac{a^2[S_t] + ab}{\sqrt{a^2[S_t]^2 + 2ab[S_t] + \left([E_t] + K_{m1}\right)^2}}$$

$$\frac{2}{ak_2 K_{m2}} \frac{d^2}{d[S_t]^2}[EP] = \frac{-4[E_t]K_{m1}}{\sqrt{a^2[S_t]^2 + 2ab[S_t] + \left([E_t] + K_{m1}\right)^2}} < 0$$

The second derivative of [*EP*] is monotonic, and approaches zero only at $[S_t] \to \infty$. Therefore, [*EP*] as a function of [S_t] is hyperbolic and shows no sigmoidal behavior.

Next we examine the function dependence of [EP] on [S] (in other words, the concentration of free [S] molecules is a control variable experimentally). One can show that the steady state solution of [EP] has the form,

$$[EP] = \frac{[S] + \alpha}{\beta[S] + \gamma} \tag{3}$$

Again its second derivative over [S] is single-signed on varying [S]. Thus the scheme shown in case a cannot have a sigmoidal-shaped [EP] ~ [S] curve.

Case a2

$$\begin{pmatrix} -k_{1f}S & k_{1r} & 0 & 0 & k_2 \\ k_{1f}S & -k_{1r} - k_{1f}^{"} & k_{1r}^{"} & 0 & 0 \\ 0 & k_{1f}^{"} & -k_{1r}^{"} - k_1 & 0 & 0 \\ 0 & 0 & k_1 & -k_{2f} & k_{2r} \\ 0 & 0 & 0 & k_{2f} & -k_{2r} - k_2 \end{pmatrix} \begin{pmatrix} E \\ ES \\ E^*S \\ EP \\ AEP \end{pmatrix} = 0 \tag{4}$$

$$\begin{aligned} E + ES + E^*S + EP + AEP &= E_t \\ S + ES + E^*S &= S_t \end{aligned} \tag{5}$$

Compared to case a, an additional intermediate step is added. After some tedious bur straightforward derivation, one obtains

$$E = \frac{-B + \sqrt{B^2 + 4AC}}{2A}$$

$$EP = \frac{\frac{k_1}{2Ak_2} K_{m2} \left(aS_t - \sqrt{a^2 S_t{}^2 + 2abS_t + b^2 + 4AC} + 2AE_t + ab \right)}{\left(\frac{k_1}{k_2} + \frac{k_1}{k_2} K_{m2} + 1 + K_{m1}^{"} \right)}$$

With

$$K_{m1} = \frac{k_{1r}}{k_{1f}}, K_{m2} = \frac{k_{2r} + k_2}{k_{2f}}, K_{m1}^{"} = \frac{k_{1r}^{"} + k_1}{k_{1f}^{"}},$$

$$A = 1 + K_{m1}^{"}$$

$$B = k_{1f} S_t \frac{k_1}{k_2} \left(1 + K_{m2} + \frac{k_2}{k_1} + \frac{k_2}{k_1} K_{M1}^{"} \right) + k_{1r} K_{m1}^{"} + k_1 - \left(1 + K_{m1}^{"} \right) E_t = aS_t + b$$

$$C = \left(k_{1r} K_{m1}^{"} + k_1 \right) E_t$$

However, similar to case a, it can be shown mathematically that the second derivative of [EP] over [S] or $[S_t]$ is single-signed. Therefore, case a2 gives no sigmoidal behavior.

B. case b is not a good candidate for switching behavior in terms of [S]

In case b, additional pathway with two S molecules bound to E is added. Here we show that no sharp switch behaviors can be obtained if the free S concentration is the control variable.

First to facilitate the following discussions (especially to make the discussions with Taylor expansions relevant), we define a reduced quantity, $s = S/S_{max}$, where S_{max} is chosen so that at which [EP] is already close to the plateau value. The two rate constants are scaled correspondingly, $\bar{k}_{1f} = k_{1f} S_{\max}, \bar{k}_{1f}^{'} = k_{1f}^{'} S_{\max}$. The governing equations are,

$$\begin{pmatrix} -\bar{k}_{1f} s & k_{1r} & 0 & k_2 & 0 \\ \bar{k}_{1f} s & -k_{1r} - k_1 - \bar{k}_{1f}^{'} s & 0 & 0 & k_{1r}^{'} \\ 0 & k_1 & -k_{2f} & k_{2r} & k_1^{'} \\ 0 & 0 & k_{2f} & -k_2 - k_{2r} & 0 \\ 0 & \bar{k}_{1f}^{'} s & 0 & 0 & -k_{1r}^{'} - k_1^{'} \end{pmatrix} \begin{pmatrix} [E] \\ [ES] \\ [EP] \\ [AEP] \\ [SES] \end{pmatrix} = 0 \qquad (6)$$

with

$$[E] + [ES] + [EP] + [AEP] + [SES] = [E_t] \qquad (7)$$

Define,

$$K_{m1} = \frac{k_{1r} + k_1}{\bar{k}_{1f}}, K_{m2} = \frac{k_{2r} + k_2}{k_{2f}}, K_{m1}^{'} = \frac{k_{1r}^{'} + k_1^{'}}{\bar{k}_{1f}^{'}}. \qquad (8)$$

Then we have,

$$[AEP] = \frac{1}{K_{m2}}[EP],\ [SES] = \frac{k_2}{K_{m2}\left(k_1 K_{m1}^{'} / s + k_1^{'}\right)}[EP],$$

$$[ES] = \frac{1}{s} K_{m1}^{'}[SES] = \frac{K_{m1}^{'} k_2}{K_{m2}\left(k_1 K_{m1}^{'} + k_1^{'} s\right)}[EP],\ [E] = \frac{K_{m1} K_{m1}^{'} k_2 + k_2 k_1^{'} s / \bar{k}_{1f}}{K_{m2}\left(k_1 K_{m1}^{'} + k_1^{'} s\right) s}[EP].$$

From Equation

$$[EP]=[E_t]\left(\frac{K_{m1}K_{m1}^{'}k_2+k_2k_1^{'}s/\bar{k}_{1f}^{'}}{K_{m2}\left(k_1K_{m1}^{'}+k_1^{'}s\right)s}+\frac{K_{m1}^{'}k_2}{K_{m2}\left(k_1K_{m1}^{'}+k_1^{'}s\right)}+1+\frac{1}{K_{m2}}+\frac{k_4s}{K_{m2}\left(k_1K_{m1}^{'}+k_1^{'}s\right)}\right)^{-1}$$

$$=[E_t]\frac{K_{m2}\left(k_1K_{m1}^{'}+k_1^{'}s\right)s}{K_{m1}K_{m1}^{'}k_2+\left(k_2k_1^{'}/\bar{k}_{1f}^{'}+K_{m1}^{'}k_2+k_1K_{m1}K_{m1}^{'}+k_1K_{m1}^{'}\right)s+\left(K_{m2}k_1^{'}+k_1^{'}+k_2\right)s^2},$$

$$=[E_t]\frac{\left(1+\frac{k_1^{'}}{k_1K_{m1}^{'}}s\right)s}{\frac{K_{m1}}{K_{m2}}\frac{k_2}{k_1}+\left(\frac{k_2k_1^{'}}{\bar{k}_{1f}^{'}k_1}\frac{1}{K_{m2}K_{m1}^{'}}+\frac{k_2+k_1}{k_1}\frac{1}{K_{m2}}+1\right)s+\left(\frac{1}{K_{m1}^{'}}\frac{k_1^{'}}{k_1}+\frac{k_1^{'}+k_2}{k_1}\frac{1}{K_{m2}K_{m1}^{'}}\right)s^2}$$

the expression for [*EP*] vs *s* is in the form,

$$[EP]=[E_t]\frac{as^2+bs}{a's^2+b's+c}$$

The coefficients are positive, and *a/a'*<1, *b/b'*<1.

$$\left([EP]/[E_t]\right)'=\frac{\left(ab'-a'b\right)s^2+2acs+bc}{\left(a's^2+b's+c\right)^2}$$

In the limit s→0,

$$\left([EP]/[E_t]\right)'\approx\frac{2(a-b')s}{c}+b/c \qquad (9)$$

To generate sigmoidal shape, the derivative should be close to 0 for small s. That is, b/c<<1, which gives relation

$$\frac{K_{m1}}{K_{m2}}\frac{k_2}{k_1} >> 1, \quad (10),$$

and (a − b') ~ 0, or

$$1 << \frac{k_2}{\bar{k}_{1f}}\frac{1}{K_{m2}} + \frac{k_2 + k_1}{k_1^{'}}\frac{K_{m1}^{'}}{K_{m2}} + \frac{k_1 K_{m1}^{'}}{k_1^{'}} + \frac{K_{m1}K_{m1}^{'}}{K_{m2}}\frac{k_2}{k_1^{'}} \quad (11).$$

Here to make it clear, 'much larger than' approximately means a factor of 10 or so.

Now let's examine the limit s→∞,

$$[EP]/[E_t] \approx \frac{1}{\left(\frac{k_1^{'} + k_2}{K_{m2}k_1^{'}} + 1\right)} \quad (12)$$

To have appreciable percentage of the enzyme exists in the form of EP in the limit s→∞, one requires $\left(1 + k_2 / k_1^{'}\right) / K_{m2} \sim< 1$. That is, the expression has the value on the order of unity or less.

It is more illuminating to examine [*EP*] at s = 1 (or [S] = [S_{max}]),

$$[EP] = [E_t]\frac{\left(1 + \frac{k_1^{'}}{k_1 K_{m1}^{'}}\right)}{\frac{K_{m1}}{K_{m2}}\frac{k_2}{k_1} + \left(\frac{k_2 k_1^{'}}{\bar{k}_{1f}k_1}\frac{1}{K_{m2}K_{m1}^{'}} + \frac{k_2 + k_1}{k_1}\frac{1}{K_{m2}} + 1\right) + \left(\frac{1}{K_{m1}^{'}}\frac{k_1^{'}}{k_1} + \frac{k_1^{'} + k_2}{k_1}\frac{1}{K_{m2}K_{m1}^{'}}\right)}$$

Because of relation (10), to obtain appreciable concentration of EP, one must have,

$$\frac{k_1^{'}}{k_1 K_{m1}^{'}} >> 1 \quad (13)$$

then,

$$[EP] \approx [E_t] \frac{1}{\frac{K_{m1}K_{m1}^{'}}{K_{m2}} \frac{k_2}{k_1^{'}} + \left(\frac{k_2}{\overline{k}_{1f}} \frac{1}{K_{m2}} + \frac{k_2 + k_1}{k_1^{'}} \frac{K_{m1}^{'}}{K_{m2}} \right) + \frac{k_1^{'} + k_2}{k_1^{'}} \frac{1}{K_{m2}} + 1} \tag{14}$$

Comparing to Equation (11), one can not satisfy the requirements that [EP] increases slowly with small s, and assumes a plateau with an appreciable value (e.g, > 0.1) for large s simultaneously. Therefore we reach the conclusion that scheme b is not a good candidate for switching. While it can show sigmoidal behavior, it may not be sharp.

On the other hand, our numerical calculations show that [EP] as a function of $[S_t]$ (calculated by Equation (6) in the main text) shows sigmoidal (sometimes toggle) switching behavior.