1 September 2000

ELSEVIER

Chemical Physics Letters 327 (2000) 38–44

CHEMICAL PHYSICS LETTERS

www.elsevier.nl/locate/cplett

Event-averaged measurements of single-molecule kinetics

Jianshu Cao *

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 14 April 2000; in final form 5 July 2000

Abstract

A modulated N-conformational-channel reactive system is used to model the recent single-molecule enzymatic experiment [Science 282 (1998) 1877]. Kinetic analysis of the model system clearly demonstrates the essential difference between ensemble-averaged bulk measurements associated with the population dynamics of full-reactions and event-averaged single-molecule measurements associated with a sequence of half-reactions. Example calculations of a two-conformational-channel system support the principal findings of the reported experiment. In particular, the observation of the focal time in the single-event distribution function and the echo signal in the two-event distribution function reveals the nature of conformational landscapes. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

The rate process in a reactive system is usually assumed to be the slowest timescale such that the reaction dynamics can be treated as a Markovian process. Then, the phenomenological exponential decay law predicts the average behavior of a bulk system, and Poisson statistics describes decay events of a finite system. However, in the presence of slow environmental fluctuations, such as in proteins and glassy systems, the simple rate prediction breaks down and the population may not decay exponentially. Yet, such environmental modulation may not be completely reflected in the phenomenological kinetics and is often smeared in bulk measurements.

Advances in single-molecule spectroscopy [1–3] allow one to observe real-time single-molecule trajectories, which consist of a chain of correlated reaction events of various lifetimes [4,5] and contain rich information of microscopic mechanisms [6–10]. For example, in a recent experiment by Lu et al. [11], enzymatic turnover events of single cholesterol oxidase molecules were monitored through the emission from the enzyme's fluorescence active site, flavin adenine dinucleotide (FAD). The statistical analysis of the single-molecule trajectories clearly demonstrated slow fluctuations in the turn-over rate of cholesterol oxidation and the dependence of the enzymatic turnovers on previous history. Xie and coworkers attributed these phenomena to slow fluctuations of protein environments and suggested a simple mechanism involving different conformational states of the enzyme. Such structural fluctuations are an essential feature of low-temperature glass systems, proteins, water and other hydrogen-

* Fax: +1-617-253-7030; e-mail: jianshu@mit.edu

0009-2614/00/$ - see front matter
PII: S0009-2614(00)00809-5

bonded systems [6–8,12]. In this Letter we present theoretical analysis of a modulated reaction model to interpret event-averaged single-molecule observations of conformational fluctuations.

2. Experiment

As illustrated in Fig. 1, the fluctuating environments are described by N discretized conformational states, with inter-conversion rate $\gamma_{a,ij}$ from the jth state to the ith state when the system is in reactant 'A', and with inter-conversion rate $\gamma_{b,ij}$ from the jth state to the ith state when the system is in product 'B'. Each conformational state defines a reversible reaction channel between reactants and products, with forward rate $k_{a,i}$ and backward rate $k_{b,i}$. This model can be viewed as a discretized version of the celebrated Agmon–Hopfield model [13]. Agmon recently extended the original model to account for the physical process in the reported enzyme turnover reaction [14]. In the fast modulation limit, the modulated reaction model reduces to a single-channel reaction with an effective rate constant and, in the slow modulation limit, it reduces to an inhomogeneous average of N channels The different conformational channels are not directly detectable since only the reactant or product is monitored by fluorescence emission. In Xie's experiment, the single FAD emission turns on in the oxidized state (reactant A) and off in the reduced state (product B), so the on/off time measures the duration that a single molecule spends in reactant/product.

We begin by studying the forward half-reaction. The master equation for the N-conformational states in reactants is written as

$$\dot{P}_a(t) = -(\Gamma_a + K_a) P_a(t), \tag{1}$$

where the vector $P_{a,i}$ is the probability of being in the ith conformational state at time t, the matrix $\Gamma_{a,ij} = \delta_{ij}\gamma_{a,ii} - \gamma_{a,ij}$, with $\gamma_{a,ii} = \sum_j \gamma_{a,ji}$, describes the conformational kinetics in reactants, and the matrix $K_{a,ij} = \delta_{ij} k_{a,i}$ describes the decay process from reactants to products. The Green's function, $G_a(t)$, is formally solved via the Laplace transformation,

$$\tilde{G}_a(z) = (zI + \Gamma_a + K_a)^{-1}, \tag{2}$$

where I is the identity matrix. Because the experimental data is averaged along single-molecule trajectories over long durations, we introduce the stationary flux $F_a = \{F_{a,i}\}$, which will be evaluated explicitly later. Now, given the vector F_a, the distribution function of on-time events is expressed as

$$f(t) = \sum K_a G_a(t) F_a = \sum_{i,j} k_{a,i} G_{a,ij}(t) F_{a,j}, \tag{3}$$

where the summation is carried over all the indices. The average on-time is given by $\langle t \rangle = \int_0^\infty f(t) t \, dt$, which then determines the rate constant $k = 1/\langle t \rangle$ used in the phenomenological kinetic description. Higher order moments of the on-time distribution can also be calculated from $\langle t^n \rangle = \int_0^\infty f(t) t^n \, dt$. Similar expressions can be found for the product state. Thus, the measurement of on-time or off-time clearly separates the forward and backward half-reactions,

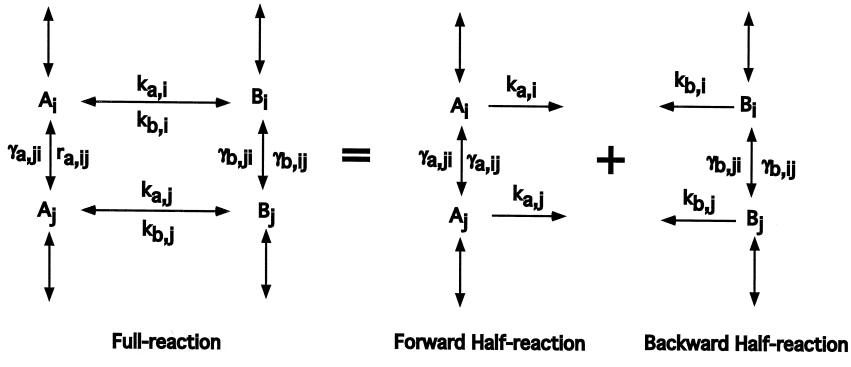

Fig. 1. The decomposition of the N conformational-channel reaction model into the forward and backward half-reactions.

and uniquely determines the statistics of single reaction events. These measurements are carried out for an ensemble of reaction events along single-molecule trajectories without keeping track of the sequence of these events.

The ingenuity of the reported experiment is to go beyond the single-event statistics to measure the correlation of a pair of on-time events along single-molecule trajectories. To couple the two half-reactions, the reactant decays to the product via rate constant k_a and the product decays to the reactant via rate constant k_b. Thus, the joint probability distribution function of two adjacent on-time events is given as

$$f(t_1,t_2) = \sum K_a G_a(t_2)\, K_b \tilde{G}_b(0)\, K_a G_a(t_1)\, F_a, \quad (4)$$

where $\tilde{G}_b(0) = \int_0^\infty G_b(t)\,\mathrm{d}t$ results from the off-time integration. In addition, the joint probability of two separate on-time events can be expressed as

$$\begin{aligned} f(t_1,t_{m+1}) = {} & \sum K_a G_a(t_{m+1})\, K_b \tilde{G}_b(0) \\ & \times \left[K_a \tilde{G}_a(0)\, K_b \tilde{G}_b(0) \right]^{m-1} \\ & \times K_a G_a(t_1)\, F_a, \end{aligned} \quad (5)$$

where m is the separation between two on-time events on single-molecule trajectories. In Ref. [11], an on-time correlation function is employed to quantitatively measure the memory effect,

$$Cor(m) = \frac{\langle t_1 t_{m+1} \rangle - \langle t \rangle^2}{\langle t^2 \rangle - \langle t \rangle^2}, \quad (6)$$

where $\langle t_1 t_{m+1} \rangle$ can be expressed in the matrix notation as

$$\begin{aligned} \langle t_1 t_{m+1} \rangle = {} & \sum K_a T_a K_b \tilde{G}_b(0) \\ & \times \left[K_a \tilde{G}_a(0)\, K_b \tilde{G}_b(0) \right]^{m-1} K_a T_a F_a, \end{aligned} \quad (7)$$

with $T_a = \int t G_a(t)\,\mathrm{d}t$. Similar expressions can be derived for off-time events and for the cross-correlation between on-time and off-time events. These quantities directly measure the non-Markovian behavior, which cannot be obtained directly in bulk experiments and must be observed along a sequence of reaction events along single-molecule trajectories.

To relate single-molecule experiments to bulk experiments, we now write the master equation for the full reaction

$$\begin{pmatrix} \dot{\rho}_a(t) \\ \dot{\rho}_b(t) \end{pmatrix} = \begin{pmatrix} -\Gamma_a - K_a & K_b \\ K_a & -\Gamma_b - K_b \end{pmatrix} \begin{pmatrix} \rho_a(t) \\ \rho_b(t) \end{pmatrix}, \quad (8)$$

where Γ_b and K_b are defined for the product, and $[\rho_a(t), \rho_b(t)]$ are the population distribution in reactants and products, respectively. The Green's function solution is given by

$$\tilde{G}(z) = (I - \Gamma - K)^{-1} = \left(I - \tilde{G}_0(z) K\right)^{-1} \tilde{G}_0(z), \quad (9)$$

where $\tilde{G}_0(z) = \begin{pmatrix} \tilde{G}_a(z) & 0 \\ 0 & \tilde{G}_b(z) \end{pmatrix}$ is the single-molecule Green's function matrix for two decoupled half-reactions, and $K = \begin{pmatrix} 0 & K_b \\ K_a & 0 \end{pmatrix}$ is the rate constant matrix between the two half-reactions. Perturbation expansion of Eq. (9) yields a series of terms in the sequence of single-molecule events, e.g.,

$$\begin{aligned} \tilde{G}_{aa}(z) = {} & \tilde{G}_a(z) + \tilde{G}_a(z)\, K_b \tilde{G}_b(z)\, K_a \tilde{G}_a(z) \\ & + \cdots, \end{aligned} \quad (10)$$

where the first term represents a single-molecule reactant without reaction, the second term represents one sojourn to product, and so on. In a sense, the time history forms a correlated random walk as a function of the number of turn-over events [15,16]. Thus, the population evolution in bulk experiments is equivalent to the summation of all the possible reaction events along single-molecule trajectories, and the equilibrium population distribution in the bulk state can be realized by time-averaging single-molecule trajectories over long durations. Consequently, single-molecule spectroscopy associated with a sequence of half-reaction events described by G_a and G_b inherently contains more detailed information about microscopic reaction mechanisms than conventional spectroscopy associated with the population dynamics described by G. Further, it is useful to distinguish two complementary measurements: the fluorescence intensity correlation function and two-

time correlation functions discussed by Schenter et al. [17], which describe the evolution of single-molecule occupancy, and the single-molecule observables discussed in Xie's original paper [11] and here, which describe the statistics of the occurrence of half-reaction events.

If conformational fluctuations are ergodic, averaging reaction events along single-molecule trajectories over long durations leads to a stationary flux, i.e., F_a or F_b used in Eqs. (3), (4), (6) and (7). To evaluate these event-averaged quantities, we first solve the equilibrium distribution of the full kinetics, $\rho = [\rho_a, \rho_b]$, by setting Eq. (8) to zero. Since the flux into reactant A from product B is $K_b \rho_b$ and the flux into product B from reactant A is $K_a \rho_a$, we have

$$\begin{pmatrix} F_a \\ F_b \end{pmatrix} = \mathcal{N} \begin{pmatrix} K_b \rho_b \\ K_a \rho_a \end{pmatrix}, \tag{11}$$

where the normalization factor is the same for the reactant and the product, i.e., $\mathcal{N}^{-1} = \sum K_b \rho_b = \sum K_a \rho_a$. Further, we can evaluate the off-time flux by averaging the on-time event in reactant A,

$$\int_0^\infty K_a G_a(t) F_a \, \mathrm{d}t = K_a (\Gamma_a + K_a)^{-1} \mathcal{N} K_b \rho_b = \mathcal{N} K_a \rho_a = F_b, \tag{12}$$

thus confirming that the occurrence flux is indeed stationary on average. Although presented in the context of the two-state model, the above analysis is general and provides a prescription for calculating the statistics of single-molecule half-reaction events.

As the two half-reactions are coupled through F_a and F_b, we discuss three possible scenarios for the forward reaction according to the relative timescale of the backward reaction.

2.1. Constant backward reaction

In the first scenario, we set the backward reaction rate to a constant, $k_{i,b} = k_b$, and make no distinction between the conformational kinetics in products and reactants, $\Gamma_a = \Gamma_b = \Gamma$. This scenario is similar to the enzymatic turnover experiment [11] by Xie and coworkers, where no fluctuations were found in the FAD oxidation half-reaction, and where both the statistics of single events and the correlation of multiple events were measured for the FAD reduction half-reaction. Though the backward half-reaction is independent of conformational fluctuations, k_b does affect the forward half-reaction through F_a and can be used as an active control parameter to reveal the underlying conformational kinetics.

2.2. Slow backward reaction

In the limit of $k_b \to 0$, the backward reaction is the rate-limiting step and is not distributed. From Eq. (11), the reactant is prepared according to the distribution of the conformation state, $F_i = \rho_{\gamma,i}$, where ρ_γ is the equilibrium solution to the modulation matrix $\Gamma \rho_\gamma = 0$. A well-studied example is ligand binding to hemoglobin under constant illumination [8]. Of particular relevance is the recent work by Wang and Wolynes [8], where non-Poisson statistics (intermittency) is explored in single-molecule reaction dynamics modulated by diffusive environments. Experimentally, after the product is formed, the system is slowly pumped back to the reactant, the cycle is repeated on single molecules, and only the single-event distribution information is collected.

2.3. Fast backward reaction

In the limit of $k_b \to \infty$, the forward reaction is rate-limiting, whereas the backward reaction is instantaneous and not sensitive to the environmental fluctuations. In this limit, the stationary flux takes the value of $F = K_a \rho_\gamma / \sum K_a \rho_\gamma$ and the macroscopic first-order rate constant becomes

$$k = \frac{1}{\langle t \rangle} = \sum K_a \rho_\gamma, \tag{13}$$

which is an inhomogeneous average of conformational channels. Hence, in this scenario, the bulk measurement does not reveal the dynamic nature of environmental fluctuations, and one has to resort to single-molecule experiments to resolve the conformational dynamics. Phenomenologically, one can treat the backward reaction as a constant source of the reactant such that the average distribution of reactive molecules is maintained at equilibrium with each conformation state. This scenario is also equivalent to the case of symmetric reactions with $K_a = K_b$,

where the on-time distribution is identical to the off-time distribution.

3. Calculations

In the rest of the Letter, we present detailed calculations of the forward half-reaction for a two-conformational-channel model, with $\gamma_{12}=\gamma_{21}=\gamma$. The Green's function in Eq. (2) is explicitly expressed as

$$G_\alpha(t)=\frac{1}{\Delta}\times\begin{pmatrix}\Delta E_c(t)-k_d E_s(t) & \gamma E_s(t)\\ \gamma E_s(t) & \Delta E_c(t)+k_d E_s(t)\end{pmatrix}, \quad (14)$$

where the subscript α refers to a for the product or b for the reactant, and the other parameters are defined accordingly. Here, $2k_d=k_{\alpha 1}-k_{\alpha 2}$, $2k_s=k_{\alpha 2}+k_{\alpha 1}$, $E_c(t)=(e^{-z_- t}+e^{-z_+ t})/2$, $E_s(t)=(e^{-z_- t}-e^{-z_+ t})/2$, and the pair of eigenvalues are $z_\pm=k_s+\gamma\pm\Delta$ with $\Delta=\sqrt{\gamma^2+k_d^2}$. The corresponding Laplace transform $\tilde{G}(z)$ in Eq. (2) and the time average T in Eq. (7) can be evaluated accordingly. Next, solving ρ from Eq. (8), we obtain the stationary flux

$$\begin{pmatrix}F_{a1}\\F_{a2}\\F_{b1}\\F_{b2}\end{pmatrix}=\mathcal{N}\begin{pmatrix}k_{a1}k_{b1}(k_{a2}+k_{b2})+\gamma k_{b1}(k_{a1}+k_{a2})\\ k_{a2}k_{b2}(k_{a1}+k_{b1})+\gamma k_{b2}(k_{a1}+k_{a2})\\ k_{a1}k_{b1}(k_{a2}+k_{b2})+\gamma k_{a1}(k_{b1}+k_{b2})\\ k_{a2}k_{b2}(k_{a1}+k_{b1})+\gamma k_{a2}(k_{b1}+k_{b2})\end{pmatrix}, \quad (15)$$

with the normalization factor $\mathcal{N}^{-1}=k_{a1}k_{b1}(k_{a2}+k_{b2})+k_{a2}k_{b2}(k_{a1}+k_{b1})+\gamma(k_{a1}+k_{a2})(k_{b1}+k_{b2})$. With the above expressions, all single-molecule observables can be calculated for the model.

As an illustrative example, we present results for a two-conformational-channel system with a constant backward rate k_b. Fig. 2 shows the on-time distribution function $f(t)$, which decays non-exponentially and depends on the choice of k_b. Interestingly, all the curves focus at t_f regardless of the backward rate. To understand this, we separate the contribution to the on-time distribution from state 1 and state 2 as

$$\begin{pmatrix}f_1(t)\\f_2(t)\end{pmatrix}=\begin{pmatrix}k_{a,1}G_{a,11}(t)+k_{a,2}G_{a,21}(t)\\ k_{a,1}G_{a,12}(t)+k_{a,2}G_{a,22}(t)\end{pmatrix}, \quad (16)$$

such that $f(t)=f_1(t)F_1+f_2(t)F_2$. Then, the focal time t_f can be determined from

$$f_1(t_f)=f_2(t_f), \quad (17)$$

which yields $\tanh(t_f\Delta)=\Delta/(k_s+\gamma)$. Thereby, at t_f, the two channels of the forward reaction become indistinguishable, and thus the forward reaction is decoupled from the backward reaction.

Perhaps, the most visual demonstration of the non-Markovian behavior in Ref. [11] is the 2-D contour plot of $f(t_1,t_{m+1})$. If the two events are not correlated, the joint distribution becomes the product of two single-event distribution functions, i.e.,

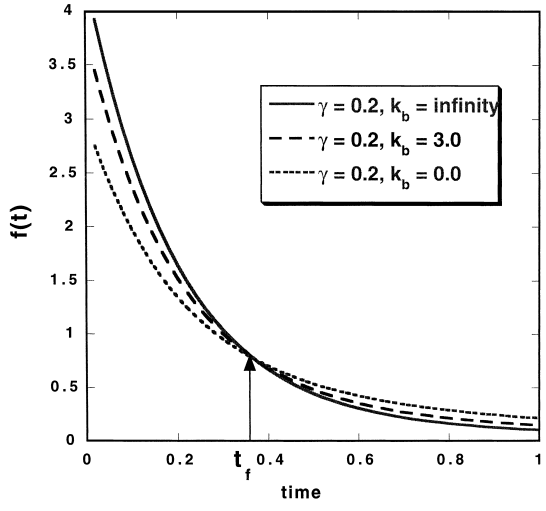

Fig. 2. A plot of the single on-time event distribution function $f(t)$ defined in Eq. (3) for the two-conformational-channel model with modulation rate $\gamma=0.2$ and with forward rate constants $k_1=5$ and $k_2=1$. For simplicity, all the time and rate variables are scaled with k_2. The three curves correspond the three scenario discussed in the text.

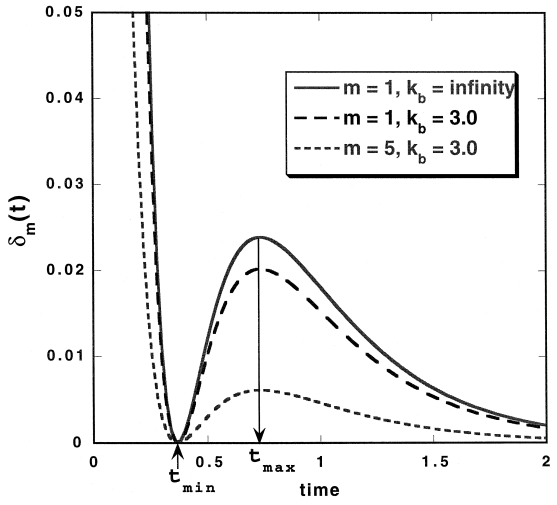

Fig. 3. A plot of the same-time difference function $\delta_m(t)$ defined in Eq. (19) for the same model as in Fig. 2 with several values of m and k_b.

$f(t_1,t_{m+1}) \to f(t_1)f(t_{m+1})$. Thus, as a probe of the memory effect, we calculate the difference distribution function

$$\delta_m(t_1,t_{m+1}) = f(t_1,t_{m+1}) - f(t_1)\,f(t_{m+1}) \qquad (18)$$

and plot the same time difference function $\delta_m(t) = \delta_m(t,t)$ in Fig. 3. As expected, the correlation decreases with time and has a larger amplitude for a smaller values of γ. In addition to the peak value at $t = 0$, we also observe an echo at a later time $t_{\max}$ and a valley with zero correlation at $t_{\min}$. The values of $t_{\min}$ and $t_{\max}$ are invariant to the backward reaction and the number of the separation between events. In fact, if normalized by the initial value at $t = 0$, all the curves in Fig. 3 are nearly the same.

To analyze this recurrent behavior, we explicitly evaluate the difference function for the third scenario ($k_b = \infty$), giving

$$\delta(t_1,t_2) = \chi_1\chi_2 H(t_1)H(t_2)\,, \qquad (19)$$

where $H(t) = 2k_d[\Delta E_c(t) - (\gamma + k_s)E_s(t)]/\Delta$ and the subscript $m = 1$ is implied. Of particular interest is the same-time difference function, $\delta(t) = \chi_1\chi_2 H(t)^2 \geq 0$, which indicates the bunching of reaction events of the same lifetime. By virtue of the extreme condition $\dot{\delta}(t) = 0$, we find the location of the valley from

$$\tanh(t_{\min}\Delta) = \frac{\Delta}{\gamma + s}\,, \qquad (20)$$

and the location of the echo from

$$\tanh(t_{\max}\Delta/2) = \frac{\Delta}{\gamma + s}\,. \qquad (21)$$

In fact, $t_{\min}$ is the same as the focal time t_f in Eq. (17), when the two channels become identical and the two adjacent events are stochastically independent. Further, $t_{\max}$ satisfies $\dot{f}_1(t) = \dot{f}_2(t)$, which implies that the difference between the two channels reaches maximum and hence two adjacent events are strongly correlated.

The same-time difference function $\delta_m(t)$ corresponds to the diagonal value of the contour plots in Ref. [11] by Xie and coworkers. Indeed, from the subtle diagonal feature of the contours, we can observe possible structures that suggest underlying molecular environments with more than two conformational channels. In reality, the relatively small amplitudes of the recurrences make it difficult to resolve peaks and valleys from experimental noise and may lead to ambiguity in identifying the number

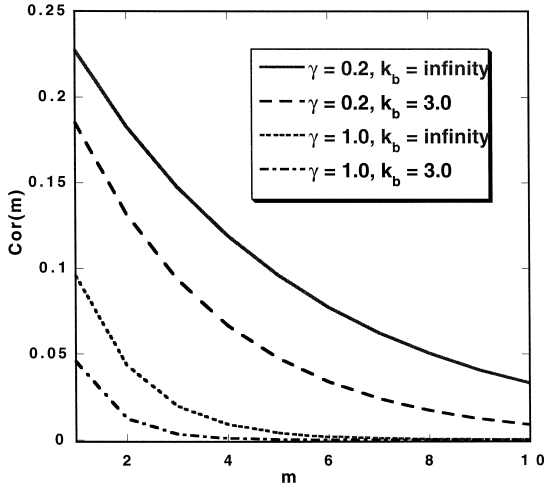

Fig. 4. A plot of the correlation function $Cor(m)$ defined in Eq. (6) for the same model as in Fig. 2 with several values of γ and k_b.

of distinctive conformational channels. Nevertheless, it is conceivable that $f(t_1,t_{m+1})$ and other related event-averaged quantities contain the essential information about the nature of conformational fluctuations.

Finally, the normalized correlation function $Cor(m)$ is plotted in Fig. 4. Since $Cor(m) = \delta_{m0}$ holds for a Markovian process, the non-vanishing value of $Cor(m)$ for $m \neq 0$ is a clear signature of the non-Markovian behavior. However, the decay of $Cor(m)$ is sensitive to the backward reaction and thus is not necessarily an accurate measure of the modulation rate.

4. Summary

The primary results of this study can be summarized as follows: (1) clarification of the differences between half and full reactions, between ensemble and event averages, and between bulk and single-molecule measurements; (2) Formulation of a general procedure for calculating the statistics of single-molecule half-reaction events; (3) explicit calculations of the two-conformational-channel model to help understand the principal features of the reported enzymatic turnover experiment; (4) identification of the focal time in the single-event distribution and the echo signal in the two-event distribution, which allow for the preliminary characterization of conformational landscapes [18]. More elaborate studies are being carried out to improve the accuracy in modeling single-molecule kinetics.

Acknowledgements

This work is supported by the MIT start-up fund. The author thanks Noam Agmon, Robert Silbey, and Sunney Xie for stimulating discussions and critical reading of the manuscript.

References

[1] T. Bache, W.E. Moerner, M. Orrit, U.P. Wild, Single-Molecule Optical Detection, Imaging and Spectroscopy, VCH, 1996.
[2] X.S. Xie, J.K. Trautman, Ann. Rev. Phys. Chem. 49 (1998) 441.
[3] W.E. Moerner, M. Orrit, Science 283 (1999) 1670.
[4] L. Edman, U. Mets, R. Rigler, Proc. Natl. Acad. Sci. USA 93 (1996) 6710.
[5] Y. Jia, A. Sytnik, L. Li, S. Vladimirov, B.S. Cooperman, R.M. Hochstrasser, Proc. Natl. Acad. Sci. 94 (1997) 7932.
[6] P. Reilly, J.L. Skinner, Phys. Rev. Lett. 71 (1993) 4257.
[7] F.L.H. Brown, R.J. Silbey, J. Chem. Phys. 108 (1998) 7434.
[8] J. Wang, P. Wolynes, Phys. Rev. Lett. 74 (1995) 4317.
[9] E. Geva, J.L. Skinner, Chem. Phys. Lett. 288 (1998) 225.
[10] A.M. Berezhkovskii, A. Szabo, G.H. Weiss, J. Chem. Phys. 110 (1999) 9145.
[11] H.P. Lu, L. Xun, X.S. Xie, Science 282 (1998) 1877.
[12] K.D. Rector, J.R. Engholm, C.W. Rella, J.R. Hill, D.D. Dlott, M.D. Fayer, J. Chem. Phys. 103 (1999) 2381.
[13] N. Agmon, J.J. Hopfield, J. Chem. Phys. 79 (1983) 2042.
[14] N. Agmon, J. Phys. Chem. B104 (2000) in press.
[15] G. Zumofen, J. Klafter, Chem. Phys. Lett. 219 (1994) 303.
[16] E. Barkai, R. Silbey, Chem. Phys. Lett. 310 (1999) 287.
[17] G.K. Schenter, H.P. Lu, X.S. Xie, J. Phys. Chem. A 103 (1999) 10499.
[18] J.N. Onuchic, J. Wang, P.G. Wolynes, Chem. Phys. 247 (1999) 89.