Observation of a Power-Law Memory Kernel for Fluctuations within a Single Protein Molecule

Wei Min,[1] Guobin Luo,[1] Binny J. Cherayil,[1,*] S. C. Kou,[2] and X. Sunney Xie[1,†]
[1]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*
[2]*Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA*
(Received 6 October 2004; published 18 May 2005)

The fluctuation of the distance between a fluorescein-tyrosine pair within a single protein complex was directly monitored in real time by photoinduced electron transfer and found to be a stationary, time-reversible, and non-Markovian Gaussian process. Within the generalized Langevin equation formalism, we experimentally determine the memory kernel $K(t)$, which is proportional to the autocorrelation function of the random fluctuating force. $K(t)$ is a power-law decay, $t^{-0.51\pm0.07}$ in a broad range of time scales (10^{-3}–10 s). Such a long-time memory effect could have implications for protein functions.

DOI: 10.1103/PhysRevLett.94.198302 PACS numbers: 82.37.-j, 02.50.-r, 05.40.-a, 87.15.He

Understanding the role of a protein's dynamic motions on its function has been a problem of long-standing interest [1]. Single-molecule experiments provide information about protein dynamics otherwise hidden in ensemble-averaged studies. Recent single-molecule investigations of a flavin oxidoreductase [2] indicate that protein conformational fluctuations occur over a broad range of time scales. Such conformational motion is closely related to the fluctuations of enzymatic rate constant [3,4]. Kou and Xie recently showed that this conformational fluctuation can be modeled by a generalized Langevin equation (GLE) [5]. Here we report a new single-molecule experiment probing equilibrium conformational fluctuation in a protein via photoinduced electron-transfer (ET). Distance fluctuations between the ET donor (D) and acceptor (A) within a protein molecule were observed over a broad range of times (10^{-3}–100 s), and their stationarity, time reversibility, and Gaussian property were proved by statistical analysis. In the GLE formalism, the autocorrelation function of the distance fluctuation was used to determine the memory kernel which turns out to be a remarkable power-law decay $K(t) \propto t^{-0.51\pm0.07}$. The broad range of time scales for conformational fluctuations at which protein reactions normally occur has implications for its biological functions, such as catalysis and allostery.

The system under study is a protein complex formed between fluorescein (FL) and monoclonal antifluorescein 4-4-20 (anti-FL). This complex is highly stable, with a small dissociation constant $K_d \sim 0.1$ nM, allowing long-time observations at the single-molecule level. Figure 1(a) shows its crystal structure, adapted from Ref. [6]. In our room temperature experiment, a single FL and anti-FL complex was first formed in solution, immobilized onto a quartz surface via the biotin-streptavidin linkage, and then repetitively excited by a 490 nm, 76 MHz, 100 fs pulse train from a frequency doubled Ti:sapphire laser. Fluorescence lifetime γ^{-1} measurements were carried out using the time-correlated single photon counting technique. The detailed experimental setup has been described previously in Ref. [2].

The fluorescence decay of a single FL molecule is monoexponential, while that of a single FL and anti-FL complex is faster and multiexponential [Fig. 1(b)]. The shorter lifetime results from photoinduced ET from the closest tyrosine residue (Tyr^{37}, donor) to FL (acceptor) [7] and is expressed by $\gamma^{-1} = (\gamma_0 + \gamma_{ET})^{-1} \approx \gamma_{ET}^{-1}$, where γ_0 denotes the fluorescence decay rate constant in

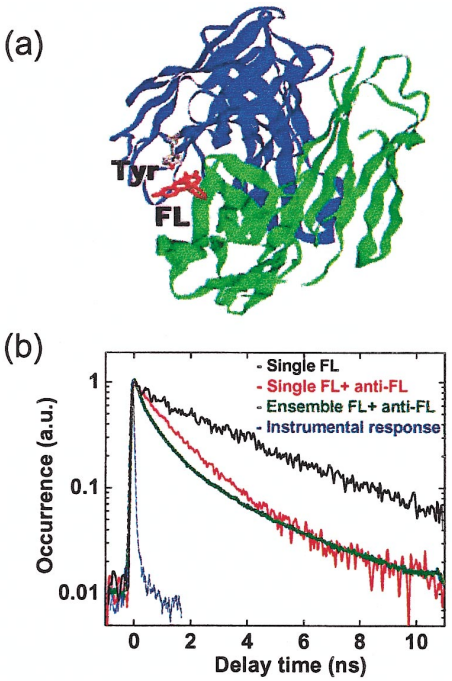

FIG. 1 (color). (a) Schematic of the structure of the FL and anti-FL complex, adapted from Ref. [6]. Tyr^{37} and FL, ET donor and acceptor, are highlighted. (b) Monoexponential fluorescence lifetime decay for a single FL molecule. Multiexponential fluorescence decay for the FL and anti-FL complex at both ensemble and single-molecule levels. The instrumental response function with 60 ps FWHM. a.u., arbitrary units.

the absence of a quencher, and γ_{ET} denotes the ET rate constant ($\gamma_{ET} \gg \gamma_0$). Figure 1(b) shows the multiexponential fluorescence decays of the FL and anti-FL complex at both the ensemble-averaged and single-molecule levels, the latter implying that fluorescence lifetime fluctuates at time scales longer than γ^{-1}.

We now discuss the origin of γ^{-1} fluctuations. According to Fermi's golden rule, the nonadiabatic ET rate γ_{ET} is $(4\pi^2/h)V^2F$, where h is Planck's constant, V is the electronic coupling between the two diabatic states DA and D^+A^-, and F is the Franck-Condon weighted density of states. $F = (4\pi k_B T\lambda)^{-1/2} \exp[-(\Delta G + \lambda)^2/4\lambda k_B T]$ at room temperature, where ΔG is the free energy difference between DA and D^+A^-, and λ is the reorganization energy [8]. For the electronically excited FL and Tyr[37], γ_{ET} is close to the maximum rate. Under this quasiactivationless condition, γ_{ET} is relatively insensitive to thermal fluctuations of ΔG. However, the electronic coupling V depends exponentially on the donor and acceptor edge-to-edge distance; $\gamma_{ET}(t)$ is therefore given by

$$\gamma_{ET}^{-1}(t) = (k_0 e^{-\beta(x_{eq}+x(t))})^{-1}, \quad (1)$$

where k_0 is a constant, x_{eq} is the mean edge-to-edge distance between FL and Tyr[37] ($\sim$3.7 Å as determined from the crystal structure [6]), $x(t)$ denotes the time-dependent distance fluctuation around x_{eq}, and $\beta \sim 1.4$ Å^{-1} for proteins [9]. Because of the sensitive exponential dependence, $x(t)$ fluctuation has a much stronger effect on $\gamma_{ET}(t)$ than the relative donor-acceptor orientation does. We thus attribute the large γ^{-1} fluctuations primarily to $x(t)$ fluctuation. Hence, from the experimentally determined $\gamma^{-1}(t)$, information about the underlying $x(t)$ fluctuation can be extracted.

By binning every 100 photons, we used the maximum likelihood estimation [10] to determine the coarse-grained $\gamma^{-1}(t)$ trajectory by assuming a single exponential fluorescence decay within each bin time. The distance fluctuation $x(t)$ trajectory is then calculated from the $\gamma^{-1}(t)$ by using Eq. (1). Figure 2(a) shows a segment of the single-molecule $x(t)$ trajectory with the corresponding probability density distribution $P(x)$. The potential of mean force $U(x)$ is then calculated from $U(x) = -k_B T \ln[P(x)]$. Figure 2(b) shows that within the thermally attainable region, $U(x)$ is well fitted by a harmonic potential $U(x) = k_B T x^2/2\theta$, where $\theta = 0.22$ Å^2. θ compares favorably with the crystal structure data, which gives average atomic-mean-square displacements of 0.12 and 0.26 Å^2 for FL and Tyr[37], respectively. This agreement further corroborates our assignment of γ^{-1} fluctuation to primarily $x(t)$ fluctuation.

The single-molecule $x(t)$ trajectory exhibits stationarity, evidenced by the same correlation functions at different time periods. The time reversibility of the $x(t)$ process is tested by evaluating $\langle x^3(0)x(t)\rangle$ and $\langle x(0)x^3(t)\rangle$. If revers-

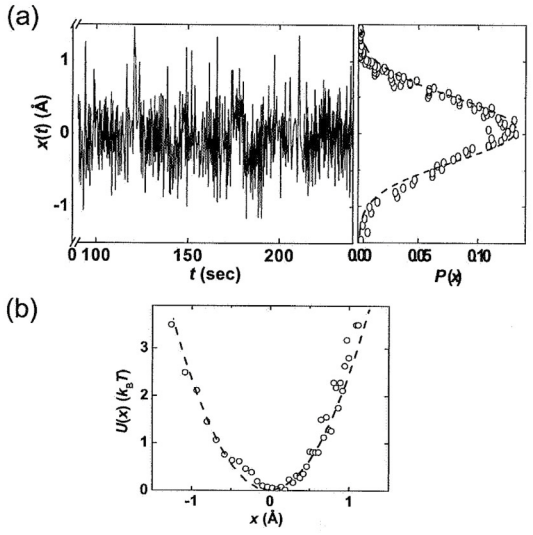

FIG. 2. (a) Segment of the single-molecule $x(t)$ trajectory with the corresponding probability density function $P(x)$ converted from $\gamma^{-1}(t)$ using Eq. (1). The dashed line is the Gaussian fit for $P(x)$. (b) Potential of mean force $U(x)$ obtained from $U(x) = -k_B T \ln[P(x)]$. The dashed line is the best fit to a harmonic potential $U(x) = k_B T x^2/2\theta$, where $\theta = 0.22$ Å^2.

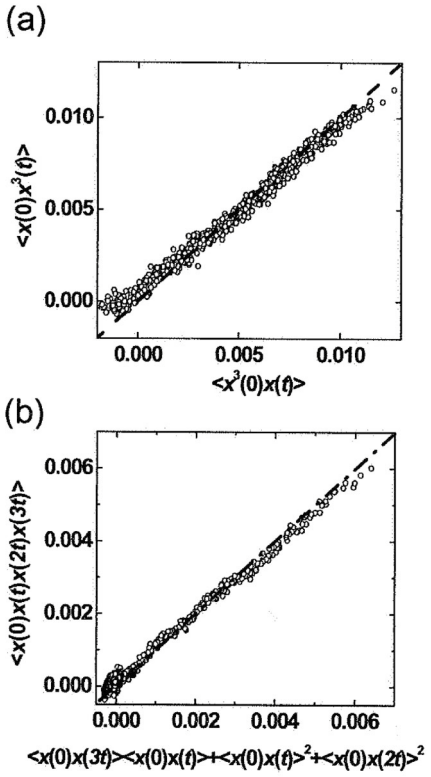

FIG. 3. (a) $\langle x^3(0)x(t)\rangle$ vs $\langle x(0)x^3(t)\rangle$ for the same single FL and anti-FL complex for various t. The diagonal dashed line is the prediction of time reversibility. (b) $\langle x(0)x(3t)\rangle\langle x(0)x(t)\rangle + \langle x(0)x(t)\rangle^2 + \langle x(0)x(2t)\rangle^2$ vs $\langle x(0)x(t)x(2t)x(3t)\rangle$ of the same complex. The diagonal dashed line is the prediction for a Gaussian process.

ibility holds, we expect

$$\langle x^3(0)x(t)\rangle = \langle x^3(-t)x(0)\rangle = \langle x(0)x^3(t)\rangle, \quad (2)$$

where the first equality is due to stationarity, and the second to reversibility. Figure 3(a) plots the experimentally determined $\langle x^3(0)x(t)\rangle$ and $\langle x(0)x^3(t)\rangle$ against each other. The diagonal line proves the time reversal symmetry.

We now examine the Gaussian property of $x(t)$. For a Gaussian process, all correlation functions higher than second order can be expressed by the second order correlation function. For example, $\langle x(0)x(t)x(2t)\rangle = 0$, and $\langle x(0)x(t)x(2t)x(3t)\rangle = \langle x(0)x(3t)\rangle\langle x(0)x(t)\rangle + \langle x(0)x(t)\rangle^2 + \langle x(0)x(2t)\rangle^2$. We calculated both $\langle x(0)x(t)x(2t)\rangle$ and $\langle x(0)x(t)x(2t)x(3t)\rangle$ from the experimental $x(t)$ trajectory and found that $\langle x(0)x(t)x(2t)\rangle$ vanishes within experimental error and that $\langle x(0)x(t)x(2t)x(3t)\rangle$ matches well with $\langle x(0)x(3t)\rangle\langle x(0)x(t)\rangle + \langle x(0)x(t)\rangle^2 + \langle x(0)x(2t)\rangle^2$ [Fig. 3(b)]. These results strongly suggest that $x(t)$ is a Gaussian process.

By virtue of the stationary and Gaussian properties of $x(t)$, $C_x(t) \equiv \langle x(t)x(0)\rangle$ is related to the autocorrelation function of fluorescence lifetime variations, $C_{\gamma^{-1}}(t)$, by

$$C_{\gamma^{-1}}(t) \equiv \frac{\langle \delta\gamma^{-1}(0)\delta\gamma^{-1}(t)\rangle}{\langle \gamma^{-1}\rangle^2} = e^{\beta^2 C_x(t)} - 1, \quad (3)$$

where $\delta\gamma^{-1}(t) = \gamma^{-1}(t) - \langle\gamma^{-1}\rangle$. $C_{\gamma^{-1}}(t)$ can be obtained with a high time resolution comparable to the reciprocal of the average photon count rate (1–2 ms), using the photon-by-photon method [11] instead of the conventional binning. Thus, $C_x(t)$ can be obtained from Eq. (3) with the same high time resolution. Figure 4 shows the averaged $C_x(t)$ of 13 molecules, and it clearly has fluctuations over a wide range of time scales. No noticeable power dependence of $C_x(t)$ in the excitation power range from 0.5 to 5 μW was observed, implying that the distance fluctuations are spontaneous rather than photoinduced.

To investigate the underlying dynamics, the fluctuation was analyzed in the framework of GLE, which can be derived from the Liouville equation using projection operators [12]. $x(t)$ is modeled as the coordinate of a fictitious particle diffusing in a potential of mean force. The GLE governing its equilibrium dynamics is

$$m\frac{d^2x(t)}{dt^2} = -\zeta\int_0^t d\tau K(t-\tau)\frac{dx(\tau)}{d\tau} - \frac{dU(x)}{dx} + F(t), \quad (4)$$

where m is the reduced mass of the particle, $U(x) = m\omega^2x^2/2$ is the harmonic potential with an angular frequency ω, ζ is the friction coefficient, $F(t)$ is the fluctuating force, and $K(t)$ is the memory kernel related to $F(t)$ by the fluctuation-dissipation theorem:

$$K(t-\tau) = (1/\zeta k_BT)\langle F(t)F(\tau)\rangle. \quad (5)$$

In the overdamped limit where acceleration can be neglected, Eq. (4) can be rewritten as

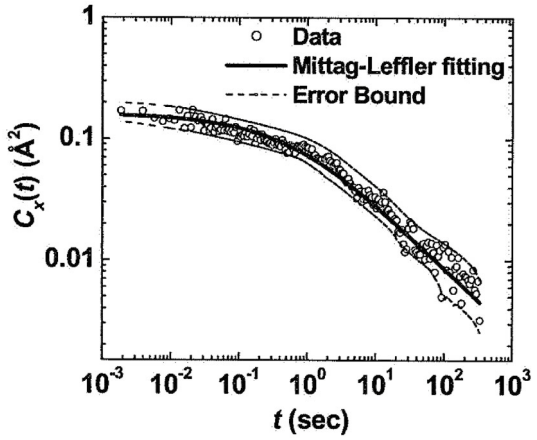

FIG. 4. Autocorrelation function of distance fluctuation $C_x(t)$ (open circles, average of 13 molecules under the same experimental condition), determined with high time resolution using Eq. (3), with $C_x(0) = k_BT/m\omega^2 = \theta = 0.22$ Å^2. The solid line is a fit to $C_x(t) = C_x(0)e^{t/t_0}\mathrm{erfc}(\sqrt{t/t_0})$ with parameter $\zeta/m\omega^2 = 0.7$ s$^{0.5}$. The error bounds (dashed line) were estimated by the method described in Ref. [17].

$$m\omega^2x(t) = -\zeta\int_0^t d\tau K(t-\tau)\frac{dx(\tau)}{d\tau} + F(t). \quad (6)$$

Equation (6) can be converted to an equation for the time correlation function $C_x(t)$ by multiplying by $x(0)$ and averaging over the initial equilibrium condition:

$$m\omega^2C_x(t) = -\zeta\int_0^t d\tau K(t-\tau)\frac{dC_x(\tau)}{d(\tau)} + \langle F(t)x(0)\rangle. \quad (7)$$

The last term $\langle F(t)x(0)\rangle = 0$ because F is orthogonal to x in the phase space [12,13]. The Laplace transform of Eq. (7) gives

$$\tilde{K}(s) = \frac{m\omega^2}{\zeta}\frac{\tilde{C}_x(s)}{C_x(0) - s\tilde{C}_x(s)}, \quad (8)$$

where $\tilde{K}(s)$ is the Laplace transform of $K(t)$. By taking the Laplace transform of $C_x(t)$ in Fig. 4 (open circles) numerically, and plugging the resulting $\tilde{C}_x(s)$ into Eq. (8) along with $C_x(0) = k_BT/m\omega^2 = \theta = 0.22$ Å^2, one solves $(\zeta/m\omega^2)\tilde{K}(s)$, which is shown in Fig. 5 after normalization. Over at least four decades of time, $\tilde{K}(s)$ exhibits a simple power-law decay, $\tilde{K}(s) \propto s^\alpha$, with $\alpha = -0.49 \pm 0.07$. Inverse Laplace transform of $\tilde{K}(s)$ gives the time domain correspondence $K(t) \propto t^{-\alpha-1} = t^{-0.51\pm0.07}$, which is remarkably simple.

The above results have implications for the nature of $F(t)$. First, since $x(t)$ is stationary, the fluctuations of $F(t)$ must likewise be stationary. Second, since GLE is a linear equation of $x(t)$, the Gaussianity of $x(t)$ requires $F(t)$ to be a Gaussian process as well. Third, the long memory behavior indicates that $F(t)$ is non-Markovian. Fourth, the power-law decay of $K(t)$ implies time scaling invariance of $\langle F(t)F(\tau)\rangle$ [Eq. (5)]. Mathematically, the only process that

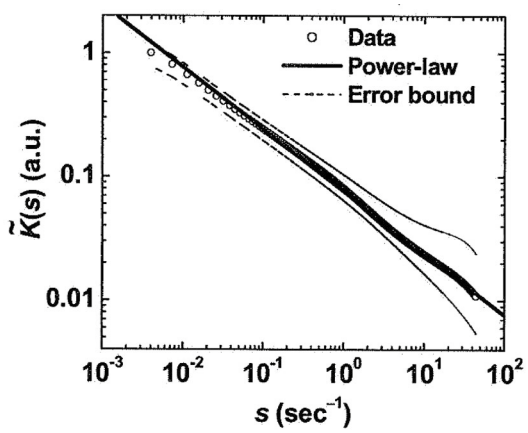

FIG. 5. Normalized $\tilde{K}(s)$ calculated from $C_x(t)$ in Fig. 4 using Eq. (8) (open circles). The full line is the fit of $s^{-0.49}$. Its inverse Laplace transform gives $K(t) \propto t^{-0.51}$. The dashed lines are error bounds carried over from the error bounds of $C_x(t)$ in Fig. 4.

simultaneously satisfies these four properties is fractional Gaussian noise (fGn) [14], which is the assumption of Ref. [5].

Following Ref. [5], $F(t)$ is assumed to be fGn defined as $F(t) = \sqrt{2k_BT\zeta}\,\frac{dB^{(H)}(t)}{dt}$, where $B^{(H)}(t)$ is the fractional Brownian motion process (Hurst coefficient $1/2 < H < 1$), with $\langle B^{(H)}(t)\rangle = 0$ and $\langle B^{(H)}(t)B^{(H)}(\tau)\rangle = (|t|^{2H} + |\tau|^{2H} - |t-\tau|^{2H})/2$. Therefore, for $t \neq \tau$,

$$K(t-\tau) = 2\left\langle \frac{dB^{(H)}(t)}{dt}\,\frac{dB^{(H)}(\tau)}{d\tau}\right\rangle = 2H(2H-1)|t-\tau|^{2H-2}. \quad (9)$$

We note that similar theoretical studies about the GLE with power-law memory kernel have appeared previously [15].

Substituting Eq. (9) into Eq. (7) and solving the resulting equation via a Laplace transform, we find $\tilde{C}_x(s) = C_x(0)/\{s + [m\omega^2/\zeta\Gamma(2H+1)]s^{2H-1}\}$, where $\Gamma(2H+1)$ is gamma function. Thus, in the time domain,

$$C_x(t) = \frac{k_BT}{m\omega^2}E_{2-2H}[-(t/t_0)^{2-2H}], \quad (10)$$

where $t_0 \equiv [m\omega^2/\zeta\Gamma(2H+1)]^{1/(2H-2)}$ is a characteristic time scale of the system, and $E_a(z) \equiv \sum_{k=0}^{\infty} z^k/\Gamma(ak+1)$ is the Mittag-Leffler function. The unit of ζ here depends on the H value. From $K(t) \propto t^{-0.51\pm0.07}$, H is determined to be 0.75 ± 0.03 using Eq. (9). With $H = 3/4$, $C_x(t) = C_x(0)e^{t/t_0}\,\mathrm{erfc}(\sqrt{t/t_0})$, where erfc is the complementary error function. Figure 4 shows the fitting of Eq. (10) using $H = 3/4$ and $\zeta/m\omega^2 = 0.7\ \mathrm{s}^{1/2}$ with the experimental $C_x(t)$. The agreement between the two is excellent. t_0 is determined to be 0.9 s.

Although nonexponential relaxations of protein conformations have been indirectly inferred from ensemble studies [16], we have directly characterized equilibrium fluctuations occurring over a broad range of time scales on which protein reactions usually take place. Our experiments demonstrate that within the experimentally accessible time ranges, the underlying fluctuating force has a simple power-law autocorrelation function. We suspect that the existence of the power-law memory kernel in proteins might be general. Investigation of its microscopic origins is under way. The fact that a protein is a complex dynamic entity with long memories on broad-range time scales leads to dynamic disorder and dispersed kinetics for biochemical reactions [1,3,4].

We acknowledge support from DOE and NSF.

*Permanent address: Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore 560012, India.
†Corresponding author.
Electronic address: xie@chemistry.harvard.edu

[1] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, Science **254**, 1598 (1991); M. Karplus, J. Phys. Chem. B **104**, 11 (2000).
[2] H. Yang *et al.*, Science **302**, 262 (2003).
[3] R. Zwanzig, Acc. Chem. Res. **23**, 148 (1990); S. Yang and J. Cao, J. Chem. Phys. **117**, 10 996 (2002); X. S. Xie, J. Chem. Phys. **117**, 11 024 (2002).
[4] H. P. Lu, L. Xun, and X. S. Xie, Science **282**, 1877 (1998); O. Flomenbom *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 2368 (2005).
[5] S. C. Kou and X. S. Xie, Phys. Rev. Lett. **93**, 180603 (2004).
[6] M. Whitlow *et al.*, Protein Eng. **8**, 749 (1995).
[7] M. Götz *et al.*, Biochemistry **41**, 4156 (2002).
[8] R. A. Marcus and N. Sutin, Biochim. Biophys. Acta **811**, 265 (1985).
[9] C. C. Moser *et al.*, Nature (London) **355**, 796 (1992).
[10] M. Kollner and J. Wolfrum, Chem. Phys. Lett. **200**, 199 (1992).
[11] H. Yang and X. S. Xie, J. Chem. Phys. **117**, 10 965 (2002).
[12] R. Zwanzig, *Lectures in Theoretical Physics* (Wiley Interscience, New York, 1961); H. Mori, Prog. Theor. Phys. **33**, 423 (1965).
[13] B. J. Berne, J. P. Boon, and S. A. Rice, J. Chem. Phys. **45**, 1086 (1966).
[14] B. Mandelbrot and J. Van Ness, SIAM Rev. **10**, 422 (1968).
[15] K. G. Wang and M. Tokuyama, Physica (Amsterdam) **265A**, 341 (1999); E. Lutz, Phys. Rev. E **64**, 051106 (2001); R. Granek and J. Klafter, Europhys. Lett. **56**, 15 (2001); R. Kupferman, J. Stat. Phys. **114**, 291 (2004).
[16] R. Austin, K. W. Beeson, L. Eisentein, and H. Frauenfielder, Biochemistry **14**, 5355 (1975); T. A. Jackson, M. Lim, and P. A. Anfinrud, Chem. Phys. **180**, 131 (1994).
[17] G. K. Schenter, H. P. Lu, and X. S. Xie, J. Phys. Chem. A **103**, 10477 (1999).