The Complex Kinetics of Protein Folding in Wide Temperature Ranges

Jin Wang
State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, People's Republic of China; and The Department of Chemistry and Department of Physics, State University of New York, Stony Brook, New York

ABSTRACT The complex protein folding kinetics in wide temperature ranges is studied through diffusive dynamics on the underlying energy landscape. The well-known kinetic chevron rollover behavior is recovered from the mean first passage time, with the U-shape dependence on temperature. The fastest folding temperature T_0 is found to be smaller than the folding transition temperature T_f. We found that the fluctuations of the kinetics through the distribution of first passage time show rather universal behavior, from high-temperature exponential Poissonian kinetics to the relatively low-temperature highly non-exponential kinetics. The transition temperature is at T_k and $T_0 < T_k < T_f$. In certain low-temperature regimes, a power law behavior at long time emerges. At very low temperatures (lower than trapping transition temperature $T < T_0/(4\sim6)$), the kinetics is an exponential Poissonian process again.

INTRODUCTION

Studying the kinetics of protein folding is the key to understanding the fundamental underlying mechanism. Levinthal (1969) posed the so-called *Levinthal paradox* 35 years ago. If protein folding proceeds with going through every possible state, then it takes cosmological time to reach the native state. In nature, protein folding finishes in millisecond to second. The recent energy landscape theory of protein folding (Bryngelson and Wolynes, 1989; Bryngelson et al., 1995; Chan and Dill, 1994; Abkevich et al., 1994; Wang et al., 1996) resolves the issue by assuming the underlying energy landscape is funneled toward the native state. Superimposed on that are the bumps and wiggles forming local traps. For folding to complete in biological timescale under physiological temperature (300 K), the slope of the funnel must be steep enough to overcome the local traps. The energy landscape theory is successful in explaining qualitatively and quantitatively many folding experiments (Bryngelson and Wolynes, 1989; Bryngelson et al., 1995; Chan and Dill, 1994; Abkevich et al., 1994; Wang et al., 1996).

Both theoretical and experimental investigations on folding and reaction kinetics show complex kinetics in different ranges of temperature (Bryngelson and Wolynes, 1989; Bryngelson et al., 1995; Chan and Dill, 1994; Abkevich et al., 1994; Wang et al., 1996; Gutin et al., 1996; Cieplak et al., 1999; Seno et al., 1998; Klimov and Thirumalai, 1998; Kaya and Chan, 2000, 2002; Itzhaki et al., 1995; Schuler et al., 2002; Lipman et al., 2003; Sabelko et al., 1999; Nguyen et al., 2003; Frauenfelder et al., 1988, 1991; Yang and Xie, 2002a,b). By varying the temperature, the underlying energy landscape structures can be probed in different levels, from the global to the local detail perspectives (Frauenfelder et al., 1988, 1991). The relationship between dynamics and functions of the biomolecules can be revealed. Although different theoretical approaches explain kinetic behavior within specific temperature ranges, the unified picture of kinetics of the whole temperature range seems still lacking. This is the purpose of the current study. Lately, the diffusive dynamics of folding is shown to have complex kinetics (Nguyen et al., 2003; Lee et al., 2003; Zhou et al., 2003; V. B. P. Leite, J. N. Onuchic, G. Stell, and J. Wang, unpublished results).

In this article, we study the kinetics in the whole temperature range. We show that the Poisson(exponential)-non-Poisson(non-exponential)-Poisson (exponential) kinetics emerge from high to low temperatures. This phenomena seems to be universal not only for protein folding, but might also exist in other biomolecular folding, biomolecular binding and reaction systems, electron transfer, viscous liquid, and glassy materials.

The current results of the study are also relevant to the single molecule studies where the mean of the observables is often unreliable due to the large statistical fluctuations (which are not smeared out by the number of molecules as in the bulk case), and cannot be used to accurately characterize the system. In general, fluctuations of the observables intrinsic for characterizing the system are obtained from the information on moments and distributions (Wang and Wolynes, 1995, 1999; Onuchic et al., 1999; Wang, 2003). The connections of the theory and simulations with the single molecule kinetic experiments can be made through the analysis of the long-time dynamic trajectories or multiple short-time runs where information on the mean, high-order moments and distributions, or histograms of the important observables can be extracted (Lu et al., 1998; Schenter et al., 1999; Moerner, 1996; Zhuang et al., 2000, 2002; Jia et al., 1999).

Submitted March 12, 2004, and accepted for publication June 28, 2004.

Address reprint requests to Jin Wang, E-mail: jinwang@sprynet.com.

0006-3495/04/10/2164/08 $2.00

doi: 10.1529/biophysj.104.042812

MATERIALS AND METHODS

In this study, we use the fraction of native conformations ρ as an order parameter to represent the folding progress. The system is assumed to be in quasiequilibrium with respect to ρ, and the states are kinetically locally connected. This is a good approximation when folding has a definite free energy barrier between non-native and native states. It might not be as good for the downhill case, where there is no barrier. The results we obtain in this article, however, seem not to be influenced qualitatively by the kinetic connectivity assumption. It has been shown that in the global kinetic connectivity case, similar kinetic behavior occurs (Saven et al., 1994; J. Wang, W. M. Huang, H. Y. Lu, and E. K. Wang, 2004, unpublished results; Wang et al., 1996). We therefore go ahead with the local kinetic connectivity—the case for diffusion on the underlying folding energy landscape in this article.

In this model there are N residues in a polypeptide chain. For each residue there are $\nu + 1$ available conformational states, one being the native state. A simplified version of the polypeptide chain energy is expressed as

$$E = -\sum \epsilon_{\mathrm{i}}(\alpha_{\mathrm{i}}) - \sum J_{\mathrm{i,i+1}}(\alpha_{\mathrm{i}}, \alpha_{\mathrm{i+1}}) - \sum K_{\mathrm{i,j}}(\alpha_{\mathrm{i}}, \alpha_{\mathrm{j}}), \quad (1)$$

where the summation indices i and j are labels for amino-acid residues, and α_{i} is the state of i^{th} residue. The three terms represent the one-body potential, two-body interactions for nearest-neighbor residues in sequence, and interactions for residues close in space but not in sequence, respectively. Due to the sequence heterogeneity, the energies and interactions can be approximated by random variables of Gaussian distributions (Derrida, 1981) with the mean biasing toward the native state (Bryngelson and Wolynes, 1989; Bryngelson et al., 1995). Along with the assumption that energies for different configurations are uncorrelated, one can easily generate an energy landscape with roughness characterized by the spreads of these probability distributions and with the mean biasing toward the native state (funneled landscape). Using a microcanonical ensemble analysis, the average free energy and thermodynamic properties of the polypeptide chain can be obtained (Derrida, 1981; Bryngelson and Wolynes, 1989; Bryngelson et al., 1995). Note that the polymer connectivity is embodied in the entropy calculations,

$$F(\rho) = N\left\{-\left(\delta\epsilon - \frac{\Delta\epsilon^2}{2T}\right)\rho - \left(\delta L - \frac{\Delta L^2}{2T}\right)\rho^2 + T\rho\log\rho + T(1-\rho)\log\frac{1-\rho}{\nu}\right\}. \quad (2)$$

$F(\rho)$ is the average free energy for the polypeptide chain. T is a scaled temperature, $\nu + 1$ is the number of conformational states of each residue, and $\delta\epsilon$ and δL are energy differences between the native and average non-native states for one- and two-body interactions (energy gap biased to native state), respectively. $\Delta\epsilon$ and ΔL are energy spreads of one- and two-body non-native interactions (roughness of the landscape). Note that the two-body energies δL and ΔL include contributions from the second and third terms in Eq. 1. The last term in F is the configurational entropy contribution.

The kinetic process along the above free-energy landscape is approximated via the use of Metropolis rate dynamics. Using continuous-time random walks, the generalized Fokker-Planck diffusion equation in the Laplace-transformed space can be obtained (Bryngelson and Wolynes, 1989; Lee et al., 2003),

$$s\tilde{G}(\rho, s) - n_{\mathrm{i}}(\rho) = \frac{\partial}{\partial\rho}\left\{D(\rho, s)\left[\tilde{G}(\rho, s)\frac{\partial}{\partial\rho}U(\rho, s) + \frac{\partial}{\partial\rho}\tilde{G}(\rho, s)\right]\right\}, \quad (3)$$

where $U(\rho, s) \equiv F(\rho)/T + \log[D(\rho, s)/D(\rho, 0)]$. In Eq. 3, s is the Laplace transform variable over time τ. $\tilde{G}(\rho, s)$ is the Laplace transform of $G(\rho, \tau)$, the probability density function. $G(\rho, \tau)d\rho$ gives the probability for a polypeptide chain to stay between ρ and $\rho + d\rho$ at time τ. The value $n_{\mathrm{i}}(\rho)$ is the initial condition for $G(\rho, \tau)$. $D(\rho, s)$ is the frequency-dependent diffusion coefficient (Bryngelson and Wolynes, 1989):

$$D(\rho, s) \equiv \left(\frac{\lambda(\rho)}{2N^2}\right)\left\langle\frac{R}{R+s}\right\rangle_{\mathrm{R}}(\rho)\Big/\left\langle\frac{1}{R+s}\right\rangle_{\mathrm{R}}(\rho), \quad (4)$$

where $\lambda(\rho) \equiv 1/\nu + (1 - 1/\nu)\rho$. The average $\langle\ldots\rangle_{\mathrm{R}}$ is taken over $P(R, \rho)$, the probability distribution function of transition rate R from one state with order parameter ρ to its neighboring states, which may have order parameters equal to $\rho - (1/\mathrm{N})$, ρ, or $\rho + (1/\mathrm{N})$. The explicit expression of $P(R, \rho)$ can be found in Bryngelson and Wolynes (1989). The boundary conditions for Eq. 3 are set as a reflecting one at $\rho = \rho_{\mathrm{i}}$ and an absorbing one at $\rho = \rho_{\mathrm{f}}$. The choice of an absorbing boundary condition at $\rho = \rho_{\mathrm{f}}$ facilitates our calculation for the first passage time and its distribution.

One can rewrite Eq. 3 in its integral-equation representation by integrating it twice over ρ:

$$\tilde{G}(\rho, s) = -\int_{\rho}^{\rho_{\mathrm{f}}} d\rho' \int_{\rho_{\mathrm{i}}}^{\rho''} d\rho''\,[s\tilde{G}(\rho'', s) - n_{\mathrm{i}}(\rho'')] \times \frac{\exp[U(\rho', s) - U(\rho, s)]}{D(\rho', s)}. \quad (5)$$

In this work we mainly study the behavior of the first passage time (FPT) for the order parameter to reach ρ_{f}. This FPT characterizes the folding time. By taking the derivatives with respect to s in Eq. 5 and taking the limit of $s = 0$, we can iteratively obtain the moments of the first passage time.

$$\langle\tau^{\mathrm{n}}\rangle = \int_0^{\infty} d\tau\,\tau^{\mathrm{n}}P_{\mathrm{FPT}}(\tau) = [n(-1)^{\mathrm{n-1}}\int_{\rho_{\mathrm{i}}}^{\rho_{\mathrm{f}}} d\rho\left(\frac{\partial}{\partial \mathrm{s}}\right)^{\mathrm{n-1}}\tilde{G}(\rho, s)]|_{\mathrm{s=0}}$$

where the $P_{\mathrm{FPT}}(\tau)$ is the distribution of the first passage time. When $n = 1$, the mean first passage time is given as

$$\bar{\tau} = \int_{\rho_{\mathrm{i}}}^{\rho_{\mathrm{f}}} d\rho \int_{\rho_{\mathrm{i}}}^{\rho} d\rho' D^{-1}(\rho)\exp\left[\frac{\mathrm{F}(\rho) - \mathrm{F}(\rho')}{\mathrm{k_B T}}\right].$$

We can also solve the integral Eq. 5 directly for $\tilde{G}(\rho, s)$, and by the observation that the distribution of the first passage time $\tilde{P}_{\mathrm{FPT}}(s) = 1 - s\tilde{\Sigma}(s)$, where $\tilde{P}_{\mathrm{FPT}}(s)$ and $\tilde{\Sigma}(s)$ are Laplace transforms of $P_{\mathrm{FPT}}(\tau)$ and $\Sigma(\tau)$ ($\Sigma(\tau) = \int_{\rho_{\mathrm{i}}}^{\rho_{\mathrm{f}}} G(\rho, \tau)d\rho$), respectively, we can obtain the information of $P_{\mathrm{FPT}}(\tau)$ by studying the behavior of $\tilde{P}_{\mathrm{FPT}}(s)$. Due to the fact that Eq. 5 is linear in $G(\rho, s)$, we can solve it with the numerical matrix-inversion technique.

RESULTS

We start the numerical calculations by setting $R_0 = 10^9\ \mathrm{s}^{-1}$, $N = 100$, and $\nu = 10$ to match the physical scales. For simplicity we assume $\delta\epsilon = \delta L$ and $\Delta\epsilon = \Delta L$. The ratio of the energy gap between the native state and the average of non-native states over the spread of non-native states, $\delta\epsilon/\Delta\epsilon$, representing the energy bias or slope toward the native state

relative to the spread (variance) or the fluctuations in energy of the polypeptide chain on the folding landscape, can be shown to be the controlling parameter in this problem. The large values of $\delta\epsilon/\Delta\epsilon$ imply a steep folding funnel and small values of $\delta\epsilon/\Delta\epsilon$ imply a rough folding landscape. We set the initial distribution to be $n_i(\rho) = \delta(\rho - \rho_i)$, where ρ_i is set to be 0.05. In our calculations we set $\rho_f = 0.95$.

The mean first passage time (MFPT) $\langle\tau\rangle$ for the folding process versus a scaled inverse temperature, T_0/T (T_0 is defined as the temperature of the minimum or optimal (fastest) mean first passage time), is plotted in Fig. 1 for several settings of the parameters $\delta\epsilon/\Delta\epsilon$, ranging from the more funneled energy landscape to the more rough energy landscape. Notice that the energy gap $\delta\epsilon$ and roughness of the landscape $\Delta\epsilon$ are dependent on both internal sequence compositions of the protein and external environments such as solvents, denaturants, pressure, etc. We have a U-shape curve for each fixed $\delta\epsilon/\Delta\epsilon$, and the MFPT reaches its minimum at temperature T_0. At high temperatures, the MFPT is large although the diffusion process itself is fast (i.e., $D(\rho, s)$ is large). This long-time folding behavior is due to the instability of the native state. The MFPT is also large at low temperature, which indicates that the polypeptide chain is trapped in low-energy non-native states. The kinetic diffusion is very slow. This is in agreement with simulation studies and the experiments (this chevron rollover phenomena was first investigated and explained by Miller et al. (1992), Socci and Onuchic (1994, 1995), Socci et al. (1996), Gutin et al. (1996), Itzhaki et al. (1995), Cieplak et al. (1999), Seno et al. (1998), Klimov and Thirumalai (1998), Kaya and Chan (2000, 2002, 2003), Chan et al. (2004), Plotkin and Wolynes (1998), Plotkin and Onuchic (2002a,b), Zhou et al. (2003), and Nguyen et al. (2003).

In the study of chevron rollover, the relationship between thermodynamics and kinetics can be revealed. Recent extensive analyses have emphasized the relative positions of the point of fastest folding and the thermodynamic folding transition (Kaya and Chan, 2003; Chan et al., 2004). The folding transition temperature T_f can be obtained from either Eq. 2 or from the crossing of the mean first passage time for unfolding time with the mean first passage time for folding. Here we plot the relative ratio of folding transition temperature T_f to the fastest folding temperature T_0, with T_f/T_0 as a function of the ratio of the energy gap to the roughness of the underlying folding energy landscape (Fig. 2). We notice that in general, the folding transition temperature T_f is higher than the fastest folding temperature T_0. The ratio T_f/T_0 slightly increases as the folding landscape is more funneled relative to the local traps. This implies that the more funneled landscape leads to larger separation of T_f with respect to T_0. This can be understood, since higher values of T_f means it is easier for folding to occur, and lower values of T_0 means it is harder to turn over to the trap regime (T_0 has the meaning of the kinetic dividing temperature of the low temperature trapping and high temperature for folding).

In view of the recent theoretical results on chevron rollovers (Kaya and Chan, 2003; Chan et al., 2004), we provide an additional figure (Fig. 3) that plots mean first passage time versus T_f/T. Notice that the horizontal axis is equal to 1 at folding temperature. As we can see the kinetic minimum of the mean first passage time occurs at the temperatures T_0 all below folding temperature T_f (since $T_0/T > 1$). This implies that at high temperatures folding is slow and the folding time is shorter as the temperature decreases (above T_0). Below folding temperature in between T_f and T_0, folding process is speeding up relative to those at the temperatures higher than T_f. This is in agreement with both experiments and theoretical investigations (Kaya and Chan, 2003; Chan et al., 2004). In fact, different sequences give different chevron rollover plots (Plotkin and Onuchic, 2002a,b; Kaya and Chan, 2003; Chan et al., 2004). As we can see clearly here, a rougher landscape (with smaller gap/roughness ratio) from a particular sequence tends to shift the kinetic curve and also the minimum or fastest folding time

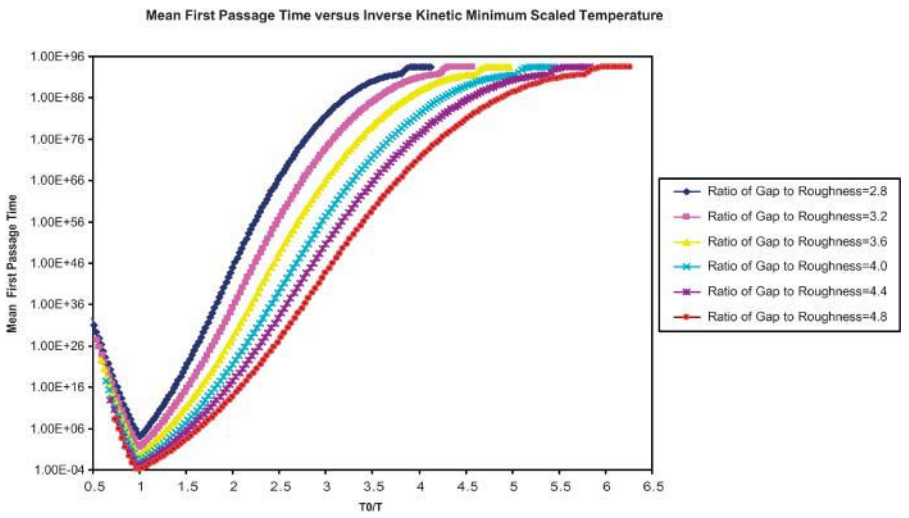

FIGURE 1 MFPT versus reduced inverse temperature T_0/T for various $\delta\epsilon/\Delta\epsilon$.

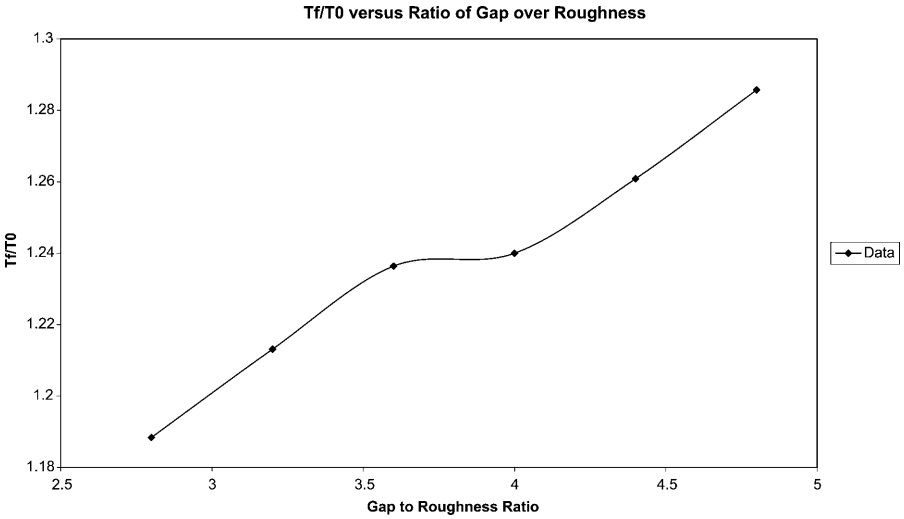

FIGURE 2 Ratio of folding temperature T_f to the kinetic folding minimum temperature T_0, T_f/T_0 versus $\delta\epsilon/\Delta\epsilon$.

toward the temperature at T_f, indicating the local traps become more and more important in determining the folding rate near T_f. Fig. 3 explicitly shows the kinetic behavior discussed earlier (Plotkin and Onuchic, 2002a,b).

We also calculate higher-order moments for the FPT distribution. In Fig. 4, we show the behavior of the reduced second moment, $\langle\tau^2\rangle/\langle\tau\rangle^2$ versus inverse kinetic minimum (fastest folding) scaled temperature. We can see that at high temperatures ($T_0/T < 1$ in this case), the ratio is a constant close to 2 (for example, the ratio is equal to 2.7 for $\delta\epsilon/\Delta\epsilon = 2.8$). Notice that for the Poissonian process, $\langle\tau^n\rangle = n!\langle\tau\rangle^n$. This implies an approximately underlying Poissonian statistics. The kinetic process can thus be shown as exponential (Wang, 2003; Lee et al., 2003; Zhou et al., 2003). On the other hand, as the temperature is lower than a specific temperature T_k, the second-order moment ratio increases rapidly. Thus the value of the temperature T_k is a marker representing the onset of the large statistical fluctuations. Note that $T_0 < T_k < T_f$. This implies that the complex kinetics occurs at a temperature T_k below folding temperature but above the fastest folding temperature. The high-order moment domination implies that the mean first passage time (MFPT) becomes unreliable for characterizing the kinetics; that the distribution of the first passage time is, in general, needed; and that it develops a fattier tail than at exponential kinetics. This implies non-Poissonian statistics and therefore non-exponential kinetics. As the temperature drops even lower ($T_0/T > 4\sim6$, the range of values is due to different landscape parameters: gap/roughness ratios), the second-order moment ratios decrease rapidly to values close to 2. This implies an approximate Poissonian process and exponential kinetics again.

In Fig. 5 we draw a graph similar to that in Fig. 4 of the reduced second moment, $\langle\tau^2\rangle/\langle\tau\rangle^2$ versus inverse folding scaled temperature T_f/T. We can see above the folding transition temperature, the second moment ratio is close to 2. This implies Poissonian exponential kinetics. Below T_f and above T_k, the second moment ratio is still close to 2. It implies

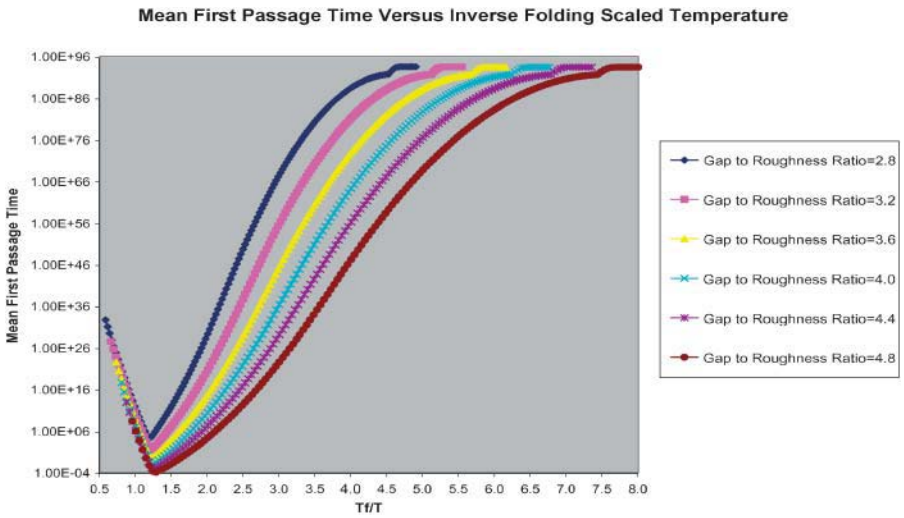

FIGURE 3 MFPT versus reduced inverse temperature T_f/T for various $\delta\epsilon/\Delta\epsilon$.

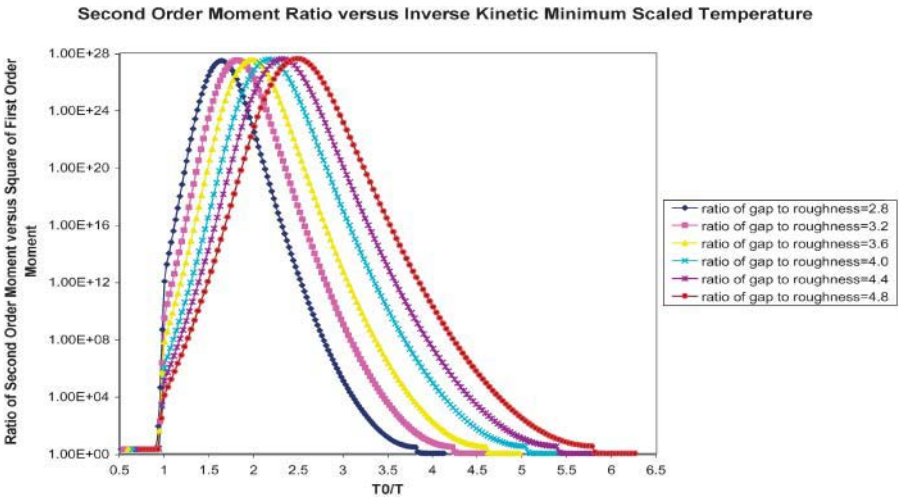

FIGURE 4 The values $\langle\tau^2\rangle/\langle\tau\rangle^2$ versus reduced inverse temperature T_0/T for various $\delta\epsilon/\Delta\epsilon$.

again Poissonian exponential kinetics. The non-exponential kinetics only emerges when the temperature is below T_k. As the temperature drops even lower ($T_f/T > 4.5 \sim 7.5$), the second-order moment ratios drop rapidly to values close to 2, implying Poissonian exponential kinetics again.

In Fig. 6, we plot the ratio of T_f/T_0, T_f/T_k, and T_k/T_0 versus gap/roughness ratio ($\delta\epsilon/\Delta\epsilon$). We found that in all the cases, $T_0 < T_k < T_f$. Since T_k is the marker for complex kinetics, this leads to exponential kinetics at temperatures $T > T_k$. The non-exponential kinetics quickly emerges at temperatures $T_0 < T < T_k$ (and continue to temperatures at (1/6~1/4) $T_0 < T < T_0$). As the gap/roughness ratio increases ($\delta\epsilon/\Delta\epsilon$), the ratio of folding temperature to kinetic transition temperature T_k (where kinetics is switched from exponential to non-exponential) increases. This indicates, as the underlying folding energy landscape is more funneled toward native state relative to the local traps, that there is a wider temperature window at $T_k < T < T_f$ for the exponential kinetics. In other words, it is relatively easier for folding (increasing T_f) and relatively harder for kinetic transition (decreasing T_k) to occur. Notice that T_k/T_0 is almost a constant. T_0 represents the onset of the rollover behavior in kinetics and T_k represents the onset of the non-exponential kinetics—the transition from a ''smoother'' appearing folding landscape to a ''rougher'' or ''bumpier'' landscape, so that basins of attraction of part of the landscape start to be partially frozen or become traps. Notice that the ratio of T_k and T_0 is almost a constant. The kinetic behavior around (T_k) is analogous to the glassy material at T_A above the glass transition temperature where partial freezing occurs (Kirkpatrick and Wolynes, 1987; Kirkpatrick et al., 1989).

In Fig. 7, the negative of the logarithm of the distribution of the first passage time in Laplace space s over 10 orders of magnitude in a Log-Log scale is plotted at an intermediate temperature ($T = T_0/2$ and $\delta\epsilon/\Delta\epsilon = 4.0$). The fact the curve looks like a straight line implies that the distribution of first passage time in Laplace space is approximately a stretched exponential ($\tilde{P}_{FPT}(s) \approx e^{-cs^u}$) which is the form of Laplace transform of the Lévy distribution in the time space. So we have

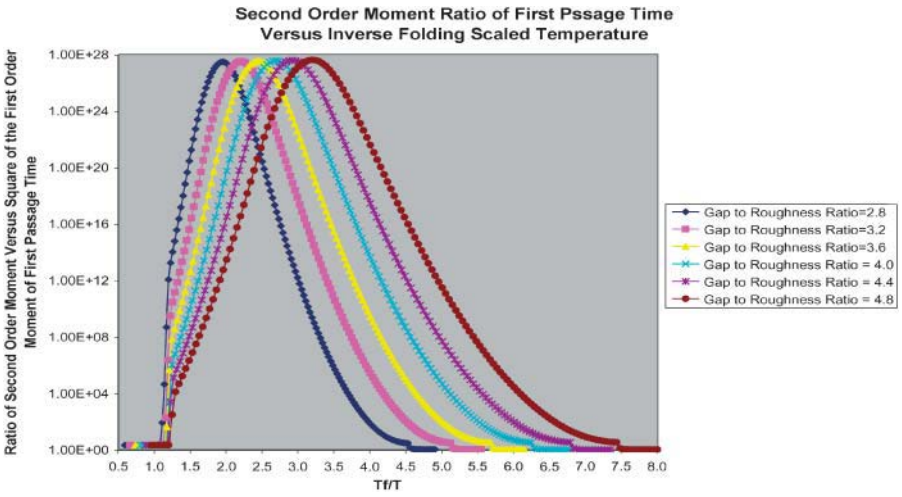

FIGURE 5 The values $\langle\tau^2\rangle/\langle\tau\rangle^2$ versus reduced inverse temperature T_f/T for various $\delta\epsilon/\Delta\epsilon$.

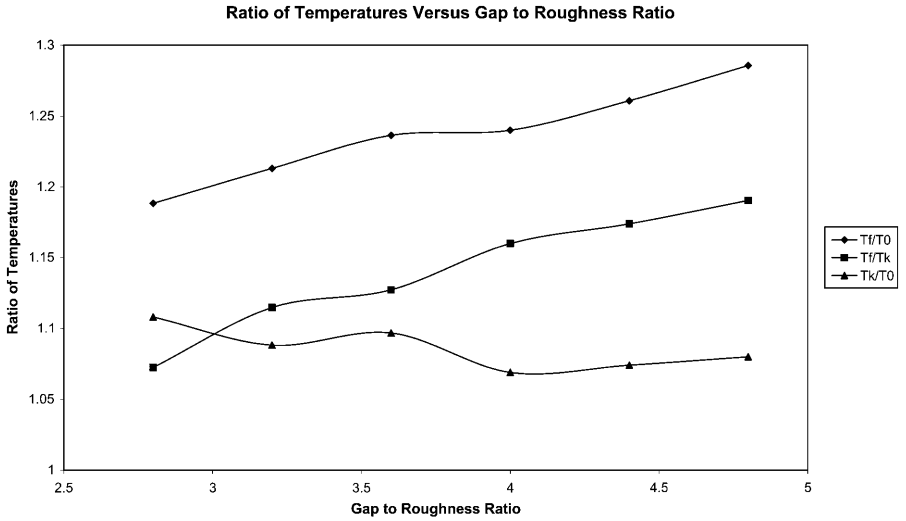

FIGURE 6 Ratio of folding temperature T_f to the kinetic folding minimum temperature T_0, T_f/T_0; ratio of folding temperature T_f to the kinetic transition temperature T_k, T_f/T_k; and ratio of kinetic transition temperature T_k to the kinetic folding minimum temperature T_0, T_k/T_0 versus $\delta\epsilon/\Delta\epsilon$.

$$P_{\mathrm{FPT}}(\tau) \approx -\frac{1}{\pi}\sum_{\mathrm{n}=1}^{\infty}\frac{(-\mathrm{c})^{\mathrm{n}}}{\tau^{\mathrm{un}+1}}\frac{\Gamma(\mathrm{un}+1)}{\Gamma(\mathrm{n}+1)}\sin(\pi u n);$$

u lies between 0 and 1. From the asymptotic property of the Lévy distribution function we learn that $P_{\mathrm{FPT}}(\tau) \sim \tau^{-(1+\mathrm{u})}$ for large τ, approaching power law distribution at long times. The power law exponent is linear with the slope of the above figure. It can be shown that power law exponent is a monotonically increasing function of the temperature below T_0, implying that the tail of the FPT distribution becomes fattier (u is smaller and the power law decay is slower) as the temperature drops lower.

DISCUSSIONS

There is a simple physical explanation of the Poisson(exponential)-power law(non-exponential)-Poisson(exponential) transition in folding kinetics: At high temperatures (higher than $T_k/T < 1$ in this case), there are multiple parallel kinetic paths leading toward the folded state and every path sees roughly similar-sized barriers (the barrier is smaller compared with kT). The highest barrier dominates and determines the kinetics since the smaller barrier can be easily overcome due to the large thermal motions. The resulting kinetics is single-exponential and the process is Poissonian. At lower temperature below T_k ($T_0/T_k < T_0/T < 4\sim6$ in this case), more and more traps become important. One can expand the Gaussian-like density of states near a frozen one, resulting in the linearization in energy of the exponential. The density of states in this temperature range thus becomes exponentially distributed. Since the first passage time is exponentially dependent on the barrier (the energy), then the distribution of first passage time follows a power law ($f(\tau)\sim 1/\tau^{1+\mathrm{T}/\mathrm{T}_0}$) for certain low temperature regimes at long times. The tail of the distribution of the first passage time becomes fattier as the temperature decreases. The rare kinetic events can play an

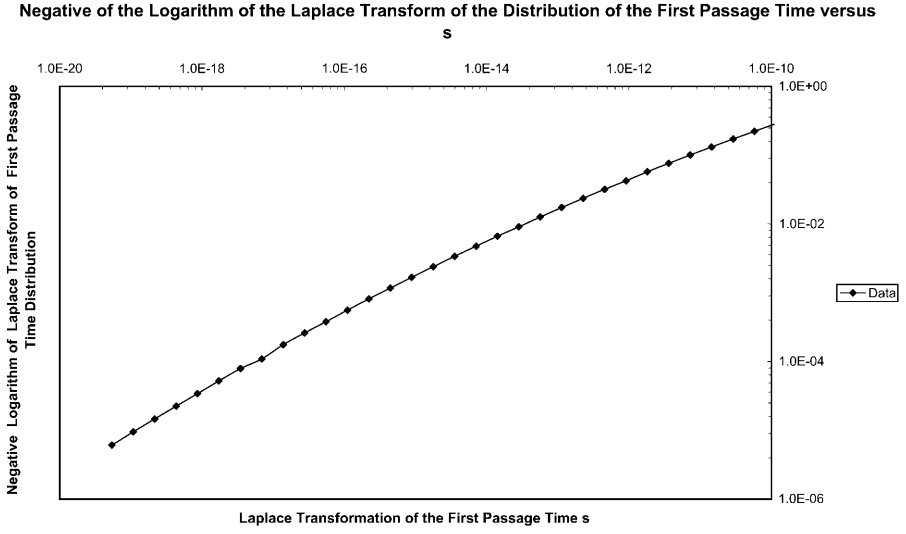

FIGURE 7 The negative of the logarithm of the distribution of first passage time is plotted in Laplace space s in a Log-Log scale.

important role. Specific kinetic path might be dominating. At very low temperature (lower then $T_0/T > 4\sim6$ in this case), only very few states are kinetically accessible. The kinetics therefore is dominated with the energy barrier in the deepest valley the system is trapped into, resulting to exponential process and Poissonian statistics again.

Four temperature scales appear in this study, all representing different scales or levels of the landscape. T_f is the folding transition temperature. The average kinetics characterized by the mean first passage time has a U-shape dependence on temperature, with the fastest time at temperature T_0. The fluctuations of the kinetics measured by the ratios of the second order moment to the square of the first-order moment has a bell-shape dependence on temperature, with the turning point at high temperature side at T_k. At very low temperatures, $T_0/T > 4\sim6$, the system is frozen (to single traps). It is important to notice that the point of the chevron rollover (the fast folding time), T_0 is related but different in value from the onset of the complex kinetics (exponential-non-exponential transition), T_k. In other words, although the turning point of the average kinetics is at T_0, the turning point of the fluctuations of the kinetics or the transition from exponential to non-exponential kinetics is at T_k (on the high temperature side). Since $T_k > T_0$ (in fact $T_f > T_k > T_0 > T_0/(4\sim6)$), the fluctuations in kinetics first sense the traps or bumps of the landscape. In other words, the fluctuations are more sensitive toward the shape of the landscape. They can be used to probe the underlying structure.

It is worth mentioning that although we focus on the study of the protein folding problem in this article, the approach we use here is very general for treating problems with barrier crossings on a complex energy landscape. In fact, several experiments on folding (Sabelko et al., 1999; Nguyen et al., 2003), binding (Frauenfelder et al., 1988, 1991), and reaction-conformational dynamics (Yang and Xie, 2002a,b) already show the existence of complex kinetics in different temperature ranges. With the rapid advances in the experiments as well as the computational power, the dynamic trajectories (long-time or multiple short-time) can be obtained and analyzed relatively easily compared with those done before (V. B. P. Leite, J. N. Onuchic, G. Stell, and J. Wang, unpublished results). It will be interesting to see the comparisons with the analytical results obtained here in the wide temperature ranges to reveal the intrinsic features and topology of the underlying energy landscape.

CONCLUSIONS

We have explored the complex kinetics of protein folding in wide temperature ranges in this study. The rather universal features emerge. The Poisson(exponential)-non-Poisson-(non-exponential)-Poisson(exponential) behavior in kinetics from high to low temperature ranges reflects the structures and topologies of the underlying energy landscapes of the protein folding at different levels. The theory and methodology outlined in this study provide a basis directly comparing and connecting with the simulations and experimental observables (V. B. P. Leite, J. N. Onuchic, G. Stell, and J. Wang, unpublished results).

I thank Prof. Peter G. Wolynes, Prof. Jose N. Onuchic, Prof. Hans Frauenfelder, Prof. Hue Sun Chan, Prof. Sunney Xie, and Prof. Martin Gruebele, Prof. George Stell and Dr. Chilun Lee for helpful discussions.

REFERENCES

Abkevich, V. I., A. M. Gutin, and E. I. Shakhnovich. 1994. Free energy landscape for protein folding kinetics: intermediates, traps, and multiple pathways in theory and lattice model simulations. *J. Chem. Phys.* 101: 6052–6062.

Bryngelson, J. D., and P. G. Wolynes. 1989. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.* 93:6902–6915.

Bryngelson, J. D., J. O. Onuchic, N. D. Socci, and P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein-folding—a synthesis. *Proteins Struct. Funct. Genet.* 21:167–195.

Chan, H. S., and K. A. Dill. 1994. Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.* 100:9238–9257.

Chan, H. S., S. Shimizu, and H. Kaya. 2004. Cooperativity principles in protein folding. *Meth. Enzymol.* 380:350–379.

Cieplak, M., T. X. Hoang, and M. S. Li. 1999. Scaling of folding properties in simple models of proteins. *Phys. Rev. Lett.* 83:1684–1687.

Derrida, B. 1981. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B.* 24:2613–2626.

Frauenfelder, H., F. Parak, and R. D. Young. 1988. Conformational substrates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* 17:451–479.

Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science*. 254:1598–1603.

Gutin, A. M., V. I. Abkevich, and E. I. Shakhnovich. 1996. Chain-length scaling of protein folding time. *Phys. Rev. Lett.* 77:5433–5436.

Itzhaki, L. S., D. E. Otzen, and A. R. Fersht. 1995. The structure of the transition state for folding of chymotrypsin inhibitor-2 analyzed by protein engineering methods—evidence for a nucleation-condensation mechanism for protein-folding. *J. Mol. Biol.* 254:260–288.

Jia, Y. W., D. S. Talaga, W. L. Lau, H. S. M. Lu, W. F. DeGrado, and R. M. Hochstrasser. 1999. Folding dynamics of single GCN-4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chem. Phys.* 247:69–83.

Kaya, H., and H. S. Chan. 2000. Energetic components of cooperative protein folding. *Phys. Rev. Lett.* 85:4823–4826.

Kaya, H., and H. S. Chan. 2002. Towards a consistent modeling of protein thermodynamic and kinetic cooperativity: how applicable is the transition state picture to folding and unfolding? *J. Mol. Biol.* 315:899–909.

Kaya, H., and H. S. Chan. 2003. Contact order dependent protein folding rates: kinetic consequences of a cooperative interplay between favorable nonlocal interactions and local conformational preferences. *Proteins.* 52:524–533.

Kirkpatrick, T. R., and P. G. Wolynes. 1987. Stable and metastable states in mean-field Potts and structural glasses. *Phys. Rev. B.* 36:8552–8564.

Kirkpatrick, T. R., D. Thirumalai, and P. G. Wolynes. 1989. Scaling concepts for the dynamics of viscous liquids near an ideal glassy state. *Phys. Rev. A.* 40:1045–1054.

Klimov, D. K., and D. Thirumalai. 1998. Linking rates of folding in lattice models of proteins with underlying thermodynamic characteristics. *J. Chem. Phys.* 109:4119–4125.

Lee, C. L., C. T. Lin, G. Stell, and J. Wang. 2003. Diffusion dynamics, moments, and distribution of first-passage time on the protein-folding

energy landscape, with applications to single molecules. *Phys. Rev. E.* 67:041905.

Lee, C. L., G. Stell, and J. Wang. 2003. First-passage time distribution and non-Markovian diffusion dynamics of protein folding. *J. Chem. Phys.* 118:959–968.

Levinthal, C. 1969. Proceedings in Mössbauer Spectroscopy in Biological Systems. P. DeBrunner, J. Tsibris, and E. Munck, editors. University of Illinois Press, Urbana, IL. p22.

Lipman, E. A., B. Schuler, O. Bakajin, and W. A. Eaton. 2003. Single-molecule measurement of protein folding kinetics. *Science.* 301:1233–1235.

Lu, H. P., L. Y. Xun, and X. S. Xie. 1998. Single-molecule enzymatic dynamics. *Science.* 282:1877–1882.

Moerner, W. E. 1996. High-resolution optical spectroscopy of single molecules in solids. *Acc. Chem. Res.* 29:563–571.

Miller, R., C. A. Danko, M. J. Fasolka, A. C. Balazs, H. S. Chan, and K. A. Dill. 1992. Folding kinetics of proteins and copolymers. *J. Chem. Phys.* 96:768–780.

Nguyen, H., M. Jäger, A. Moretto, M. Gruebele, and J. W. Kelly. 2003. Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc. Natl. Acad. Sci. USA.* 100:3948–3953.

Onuchic, J. N., J. Wang, and P. G. Wolynes. 1999. Analyzing single molecule trajectories on complex energy landscapes using replica correlation functions. *Chem. Phys.* 247:175–184.

Plotkin, S. S., and P. G. Wolynes. 1998. Non-Markovian configurational diffusion and reaction coordinates for protein folding. *Phys. Rev. Lett.* 80:5015–5018.

Plotkin, S. S., and J. N. Onuchic. 2002a. Understanding protein folding with energy landscape theory. I. Basic concepts. *Quar. Rev. Biophys.* 35:111–167.

Plotkin, S. S., and J. N. Onuchic. 2002b. Understanding protein folding with energy landscape theory. II. Quantitative aspects. *Quar. Rev. Biophys.* 35:205–286.

Sabelko, J., J. Ervin, and M. Gruebele. 1999. Stretching lattice models of protein folding. *Proc. Natl. Acad. Sci. USA.* 96:6031–6036.

Saven, J. G., J. Wang, and P. G. Wolynes. 1994. Kinetics of protein folding: the dynamics of globally connected rough energy landscapes with biases. *J. Chem. Phys.* 101:11037–11043.

Seno, F., C. Micheletti, A. Maritan, and J. R. Banavar. 1998. Variational approach to protein design and extraction of interaction potentials. *Phys. Rev. Lett.* 81:2172–2175.

Schenter, G. K., H. P. Lu, and X. S. Xie. 1999. Statistical analyses and theoretical models of single-molecule enzymatic dynamics. *J. Phys. Chem. A.* 103:10477–10488.

Schuler, B., E. A. Lipman, and W. A. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 419:743–747.

Socci, N. D., and J. N. Onuchic. 1994. Folding kinetics of proteinlike heteropolymers. *J. Chem. Phys.* 101:1519–1528.

Socci, N. D., and J. N. Onuchic. 1995. Kinetic and thermodynamic analysis of proteinlike heteropolymers: Monte Carlo histogram technique. *J. Chem. Phys.* 103:4732–4744.

Socci, N. D., J. N. Onuchic, and P. G. Wolynes. 1996. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* 104:5860–5868.

Wang, J., and P. G. Wolynes. 1995. Intermittency of single molecule reaction dynamics in fluctuating environments. *Phys. Rev. Lett.* 74:4317–4320.

Wang, J., J. Onuchic, and P. Wolynes. 1996a. Statistics of kinetic pathways on biased rough energy landscapes with applications to protein folding. *Phys. Rev. Lett.* 76:4861–4864.

Wang, J., J. G. Saven, and P. G. Wolynes. 1996b. Kinetics in a globally connected, correlated random energy model. *J. Chem. Phys.* 105:11276–11284.

Wang, J., and P. G. Wolynes. 1999. Intermittency of activated events in single molecules: the reaction diffusion description. *J. Chem. Phys.* 110:4812–4819.

Wang, J. 2003. Statistics, pathways and dynamics of single molecule protein folding. *J. Chem. Phys.* 118:952–958.

Yang, H., and X. S. Xie. 2002a. Statistical approaches for probing single-molecule dynamics photon by photon. *Chem. Phys.* 284:423–437.

Yang, H., and X. S. Xie. 2002b. Probing single-molecule dynamics photon-by-photon. *J. Chem. Phys.* 117:10965–10979.

Zhou, Y., C. Zhang, G. Stell, and J. Wang. 2003. Temperature dependence of the distribution of the first passage time: results from discontinuous molecular dynamics simulations of an all-atom model of the second-hairpin fragment of protein G. *J. Am. Chem. Soc.* 125:6300–6305.

Zhuang, X. W., L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu. 2000. A single-molecule study of RNA catalysis and folding. *Science.* 288:2048–2051.

Zhuang, X. W., H. Kim, M. J. B. Pereira, H. P. Babcock, N. G. Walter, and S. Chu. 2002. Correlating structural dynamics and function in single ribozyme molecules. *Science.* 296:1473–1476.