Diffusive Model of Protein Folding Dynamics with Kramers Turnover in Rate

Robert B. Best* and Gerhard Hummer†

Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

(Received 4 November 2005; published 7 June 2006)

We study the folding kinetics of a three-helix bundle protein using a coarse polymer model. The folding dynamics can be accurately represented by one-dimensional diffusion along a reaction coordinate selected to capture the transition state. By varying the solvent friction, we show that position-dependent diffusion coefficients are determined by microscopic transitions on a rough energy landscape. A maximum in the folding rate at intermediate friction is explained by ''Kramers turnover'' in these microscopic dynamics that modulates the rate via the diffusion coefficient; overall folding remains diffusive even close to zero friction. For water friction, we find that the ''attempt frequency'' (or ''speed limit'') in a Kramers model of folding is about 2 μs^{-1}, with an activation barrier of about $2k_BT$, and a folding transition path duration of $\approx$100 ns, 2 orders of magnitude less than the folding time of $\approx$10 μs.

DOI: 10.1103/PhysRevLett.96.228104 PACS numbers: 87.15.Cc, 87.15.He

Theory and simulation suggest that the essential features of protein folding can be captured by diffusion along a well-chosen reaction coordinate, such as the fraction of native amino acid contacts [1]; this enormous reduction of complexity can be explained by a globally funneled energy landscape [2–5]. Diffusive models of dynamics have been successful in predicting protein folding rates [6] and mechanisms. In addition, recent experimental studies of fast-folding proteins have been interpreted within the framework of diffusion. For example, diffusion-limited chain collapse and contact formation can be used to set a limit for the maximum possible rate of folding [7]. For downhill folding scenarios, where the free energy barrier is vanishingly small, a diffusive model is essential [8,9]. A low-dimensional diffusive description has also been useful in other biophysical processes [10,11].

While theories of folding are often formulated in terms of a reaction coordinate, detailed atomistic simulations and the folding of real proteins occur in a high dimensional space. Is it possible to embed such intricate dynamics into a low-dimensional representation, which would aid in the interpretation of the dynamics and permit direct comparison with theory? We address this question, by studying the folding of a Gō-like model [12] for the 47-residue three-helix bundle protein prb_{7-53} [13], in which each amino acid residue is represented by a single particle [14]. This is a fast-folding protein which predominantly populates only two states at equilibrium (folded and unfolded). We have shown that the fraction of native contacts (Q) is a nearly optimal coordinate for identifying transition states [15] for this protein. Here, we investigate whether this good folding coordinate is also dynamically relevant, assessed by how well a Markovian diffusion model captures the projected (non-Markovian) dynamics along Q.

Langevin simulations of the model protein at its folding temperature ($T_f = 292$ K) include many folding transitions, as monitored by the projection onto Q [Figs. 1(a) and 1(b)]. To analyze the trajectories quantitatively, the dynamics is modeled as Markovian diffusion along a one-dimensional coordinate (Q in this case), described by a discretized Smoluchowski equation $\partial_t p = \partial_Q(D(Q) \times \exp[-\beta F(Q)]\partial_Q\{\exp[\beta F(Q)]p\})$ [16], where $p = p(Q, t)$ is the probability density of Q at time t, β is the inverse temperature, and $D(Q)$ and $F(Q)$ are position-dependent diffusion coefficients and free energies, respectively. A transition matrix is constructed by counting ''transitions'' between intervals along Q after a lag time Δt, to allow for averaging over fast, non-Markovian processes. A Bayesian approach is used to obtain the most probable $D(Q)$ and $F(Q)$, given the trajectory $Q(t)$, by comparing the transition matrices computed from the simulation with transition matrices obtained by solving the Smoluchowski equation parameterized by $D(Q)$ and $F(Q)$ [17] [Figs. 1(c) and 1(d)]. For lag times greater than 2 ns, the folding rate back calculated from the lowest nonzero eigenvalue of the diffusion operator converges within statistical error to that obtained directly from the simulation [inset to Fig. 1(c)], indicating that a diffusive model fits the data well [1]. Deviations at short lag times Δt result from fast processes not captured by Q, leading to non-Markovian dynamics for $\Delta t \lesssim 1$ ns. In practice, the fitted $F(Q)$ are essentially identical to those obtained from histograms.

What determines the apparent diffusion coefficient $D(Q)$? Diffusive folding models commonly assume a uniform diffusion coefficient that is proportional to the reciprocal of the solvent friction. However, we find a markedly position-dependent $D(Q)$ with a nonmonotonic dependence on solvent friction (both from Bayesian and harmonic oscillator estimates; Fig. 1). Whereas the position dependence can be removed by a simple variable transformation, the friction dependence is nontrivial, and points to the microscopic events determining $D(Q)$.

To explore the connection between microscopic and global dynamics (i.e., folding), we study the kinetics of folding over a broad range of friction coefficients. Folding rates were calculated from equilibrium trajectories and complemented using rates from transition path sampling (TPS) [15,18,19] (Fig. 2). In this version of TPS, trajectory

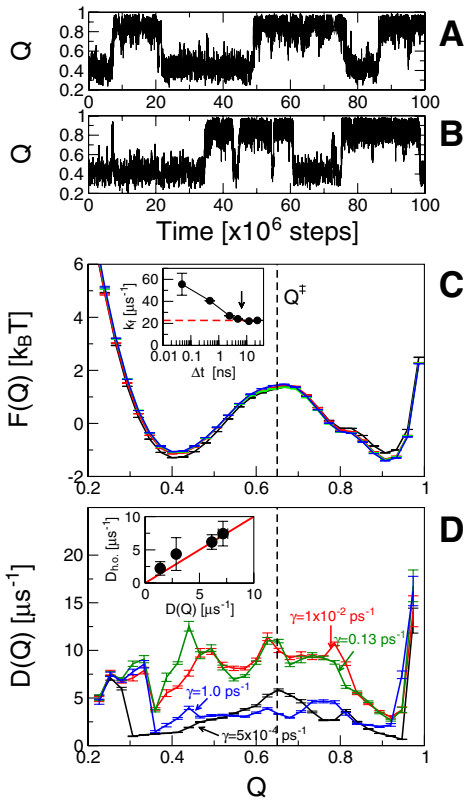

FIG. 1 (color online). Diffusive folding dynamics of a three-helix bundle protein. (a) and (b) show portions of trajectories projected onto Q for Langevin friction coefficients γ of 5.0×10^{-4} and 1.0 ps^{-1}, respectively. From the equilibrium trajectories, a Bayesian approach was used to estimate (c) the free energy surface $F(Q)$ and (d) position-dependent diffusion coefficients $D(Q)$ from the trajectories $Q(t)$ [17]. Trajectories with Langevin friction of $\gamma = 5.0 \times 10^{-4}$ (black curve), 1.0×10^{-2} (red curve), 0.13 (green curve), and 1.0 (blue curve) ps^{-1} were analyzed. The position of the transition state (as defined in the text; $Q^{\ddagger} = 0.65$) is indicated by the vertical broken line in (c) and (d) [this is identical to the top of the barrier in $F(Q)$ within the resolution of the analysis]. Inset to (c) shows the convergence of folding rate ($\gamma = 0.13$ ps^{-1}) with lag time Δt used to construct the diffusion equation; the arrow indicates the lag used to derive the $F(Q)$ and $D(Q)$ shown. Inset to (d) shows diffusion coefficients calculated using the harmonic oscillator (h.o.) approximation compared with those estimated from the diffusive model. The h.o. diffusion coefficients were obtained from simulations with a harmonic potential biasing Q toward 0.4 using the approximation, $D_{\mathrm{h.o.}} = \mathrm{Var}(Q)/\tau_Q$, where $\mathrm{Var}(Q)$ and τ_Q are, respectively, the variance and correlation time for fluctuations of Q [1].

pairs with positive and negative velocities (i.e., forward and backward in time) are initiated from equilibrium conformations on an arbitrary dividing surface (here: $Q^{\ddagger} = 0.65$) and appropriately weighted [15,19]. TPS rates were calculated using [15,19]

$$2c_{\mathrm{f}}k_{\mathrm{u}} = 2c_{\mathrm{u}}k_{\mathrm{f}} \approx p_{\mathrm{eq}}(Q^{\ddagger})\left\langle \theta_{\mathrm{TP}}\left(\sum_k |\dot{Q}_k|^{-1}\right)^{-1}\right\rangle_{\ddagger}, \quad (1)$$

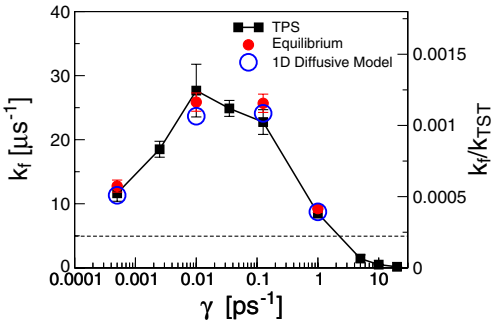

FIG. 2 (color online). Dependence of folding rates on Langevin friction. Rate coefficients were calculated from TPS (black squares, solid line), from long equilibrium simulations (red circles) and from the diffusive model (open blue circles). The right hand axis indicates the ratio of the rate to that obtained from TST, estimated as $k_{\mathrm{TST}} = p_{\mathrm{eq}}(Q^{\ddagger})\langle|\dot{Q}|\rangle_{\ddagger}$ [20]: $p_{\mathrm{eq}}(Q^{\ddagger})$ and $\langle|\dot{Q}|\rangle_{\ddagger}$ are, respectively, the equilibrium probability density and the average time derivative of Q (fraction of native contacts) at the barrier. The broken line represents the rate at zero friction, obtained from TPS with Hamiltonian dynamics.

with $\theta_{\mathrm{TP}} = 1$ if the trajectory pair forms a transition path (i.e., crosses between $Q = 0.4$ and $Q = 0.9$), and zero otherwise. $p_{\mathrm{eq}}(Q)$ is the equilibrium probability density of Q, and c_{u} and c_{f} are the mole fractions of unfolded and folded protein, respectively. The sum is over the velocities of crossing the $Q = Q^{\ddagger}$ surface. Both the form of Eq. (1) and its implementation resemble reactive flux methods [20]. Simulations were run using Langevin dynamics with the Brooks-Brünger-Karplus [21] algorithm in CHARMM [22], with a time step of 12 fs. The same rates were obtained with a shorter time step (5 fs), or a different stochastic integration algorithm [23,24] (data not shown).

As shown in Fig. 2, rates determined from equilibrium and TPS simulations as well as one-dimensional diffusion are fully consistent and show a distinct maximum at a friction coefficient of $\gamma \approx 0.01$–0.1 ps^{-1}. Only at frictions $\gamma \gtrsim 1$ ps^{-1}, do the folding rates follow the γ^{-1} dependence expected from Kramers theory for one-dimensional diffusion over a single free energy barrier [20,25]. The folding rate at zero friction (i.e., Hamiltonian dynamics), 4.93 $\mu\mathrm{s}^{-1}$, is $\approx 20\%$ of the maximum rate. While slowing, or "turnover" of the rate is indeed predicted in the one-dimensional Kramers model [19,25], it may not at first be expected for a protein, which has many more degrees of freedom. The reduction in the rate at low friction in the 1D Kramers problem is caused by the very weak exchange of energy with the heat bath. As a result, the dynamics become inertia dominated: either the system has insufficient energy to cross the barrier, or if it crosses, is unable to release the energy in order to remain in the product state. In the present case, however, the fit of the diffusive model and the appearance of the trajectories (Fig. 1) indicate that the trajectories remain diffusive at all values of friction studied, without any indication of the characteristic rapid recrossing events of low-friction Kramers theory. Consis-

tent with diffusive folding, even the maximum folding rate is 3 orders of magnitude slower than the transition state theory (TST) rate.

The Kramers-like turnover in both folding rates k_f and position-dependent diffusion coefficients $D(Q)$ suggests that low-dimensional microscopic transitions on a rough energy landscape must be important in determining $D(Q)$, and in turn k_f. A simple analogy is the variation of diffusion coefficients with friction in a jump-diffusion model [26] in which the diffusion in the global coordinate arises from hopping between many local minima. At long times, motion on such a landscape appears diffusive, with a diffusion coefficient proportional to the hopping rate between minima, and it is this rate for crossing microscopic barriers that varies in a Kramers-like way with friction, γ.

Here, backbone dihedral angle transitions, together with amino acid contact formation and breaking, are the dominant microscopic barrier-crossing events. To quantify the effect of friction on the dihedral dynamics, we study rotations about the bond between residues 19 and 20 in a four-residue fragment (residues 18–21) and a 12-residue fragment (residues 14–25) in which all attractive nonbonded interactions are turned off. Relaxation times were measured by running Langevin dynamics simulations and calculating the correlation function $\langle h(t)h(0)\rangle$ of the indicator function $h(t)$ defined as zero in one rotameric state and one in the other. Rate coefficients estimated from exponential fits to these correlation functions are given in Fig. 3. For the minimal four-residue fragment, the relaxation time exhibits a strong Kramers-like turnover with friction. In fact, it can be well approximated by a Kramers model for a double-well potential, since this dihedral angle has only two minima of approximately equal energy, and the flux is primarily over one barrier, because the other is $\approx 1.9 k_B T$ higher in energy (inset to Fig. 3). Figure 3 also compares the simulation data with the Mel'nikov and Meshkov [27] formula for the Kramers rate coefficients, using parameters derived from the simulations, without any fitting to the rates (see legend to Fig. 3 for details); the agreement with the Kramers model is good considering the approximations involved. In contrast to the protein, dihedral transitions at low friction do exhibit recrossings of ballistic character [Fig. 3(b)]. However, this fragment is clearly not a completely satisfactory model for torsional barrier crossing in the protein, since the turnover occurs at higher friction (≈ 1.0 ps^{-1}) than it does in the diffusion coefficients $D(Q)$ and folding rates for the protein (≈ 0.01–0.1 ps^{-1}); furthermore, the rates in the fragment approach zero in the limit of zero friction, unlike in the protein.

The larger 12-residue fragment shows more "protein-like" character: the "rate" is no longer zero for $\gamma = 0$, k/k_{TST} is smaller, the turnover occurs at lower friction and the trajectories are more diffusive. These deviations from Kramers-like behavior likely reflect "internal friction" arising from nonlinear couplings between the additional degrees of freedom of the peptide (8 extra residues) [25,28,29] and non-Markovian effects resulting from comparable time scales of isomerization of adjacent dihedral angles [30,31]. Nonetheless, there is a good correspondence between the turnover in the rate of jumping between dihedral isomers observed for the longer fragment (Fig. 3) and that observed for the folding rates and diffusion coefficients at the barrier (Figs. 1 and 2).

Ours is not the first computational investigation of the effect of friction on folding kinetics. Zagrovic and Pande [32] found a transition from an inverse dependence of rate on friction to a power law dependence at around 1/10 the viscosity of water, but no low viscosity turnover. Most notably, a study by Klimov and Thirumalai using coarse folding models [33] found strong turnover, with rates at the lowest friction being over an order of magnitude smaller than at the maximum. However, the turnover was related to crossing the folding barrier. Here, we find that for our

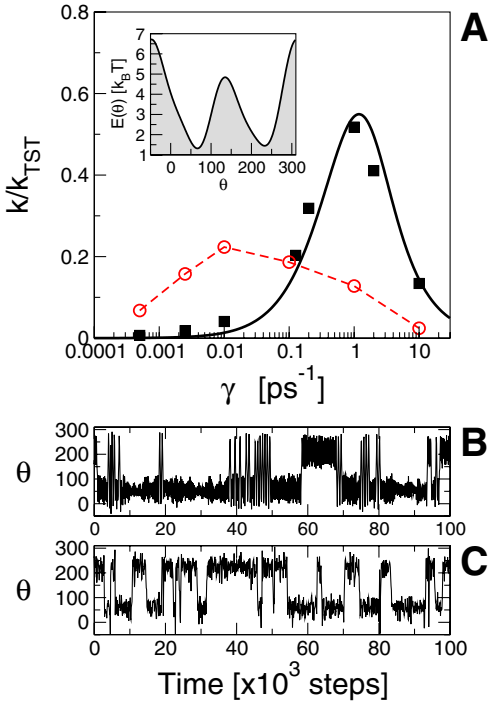

FIG. 3 (color online). (a) Rates for isomerization of the central dihedral angle for a minimal 4-residue fragment (residues 18–21, filled black squares) and a 12-residue fragment (residues 14–25, open red circles). Since the torsional potential (inset) is approximately a double well, rates were calculated as $k \approx 1/2\tau_c$ from the correlation time τ_c (TST rates were obtained as in Fig. 2). The solid black line shows the Mel'nikov-Meshkov formula [27] for the Kramers rates obtained using parameters derived from the 4-residue fragment simulations. An effective mass m_{eff} was estimated from the simulation data using the equipartition theorem for kinetic energy. The barrier height ($3.3 k_B T$) was taken from the height of the torsional potential $V(\theta)$, and from its curvature at the barrier $\theta^{\ddagger}$, $V''(\theta^{\ddagger})$, a barrier frequency $\omega_b = (V''(\theta^{\ddagger})/m_{\text{eff}})^{1/2}$ of 1.47 ps^{-1} was obtained; we assumed the well frequency $\omega_0 = \omega_b$. In (b) and (c) are shown segments of trajectories for the four-residue fragment at $\gamma = 5 \times 10^{-4}$ ps^{-1} and $\gamma = 1$ ps^{-1}, respectively.

system the dynamics appears diffusive over the whole range of friction coefficients considered. We show that the slowdown at low friction is caused by a decrease in the effective diffusion coefficient which we attribute to turnover in the transitions over the microscopic barriers on a rough energy landscape.

A second finding of the previous work [33] was that for sequences with multiple folding pathways, changing from low to intermediate friction can change the balance (''kinetic partitioning'') between these pathways. For our system, there does not appear to be a significant change of mechanism with viscosity. We find that the position of the transition state on Q (defined as that Q which has the highest probability that trajectories passing through it are reactive [15,19]) does not vary appreciably with friction. In this sense, our model is probably more like the fast-folding sequences of Klimov and Thirumalai, where no change of mechanism was found [33]. On the other hand, the variation in the position-dependence of the diffusion coefficients $D(Q)$ indicates that there is a change in the dynamics as a function of the friction.

By scaling the rate calculated at a friction of 10 ps^{-1} to a value commonly used to mimic water, 50 ps^{-1}, we obtain a folding time of $\tau_f = 1/k_f$ of 10.1 μs, in surprisingly good agreement with experiment (6 μs) for the wild-type protein. The average transition path length is approximately 1.5 ns for low to intermediate friction, increasing approximately linearly with friction for $\gamma > 1$ ps^{-1}. From the average path length of 20 ns at $\gamma = 10$ ps^{-1}, we therefore estimate that the transition path duration should be approximately 100 ns at $\gamma = 50$ ps^{-1}. Our model can also be used to estimate the ''preexponential factor'' for protein folding: assuming equal curvature of the wells and barrier, and high friction dynamics, the folding time is given by $\tau_f \approx 2\pi\tau_0 \exp(\Delta G^{\ddagger}/k_B T)$, where τ_0 is the reconfiguration time in the unfolded well. Combining the free energy barrier to folding for our model, $\Delta G^{\ddagger} \approx 2.5 k_B T$, with the experimental folding time in water of 6 μs, we estimate the preexponential factor $2\pi\tau_0 \approx 0.5$ μs, below the experimentally determined upper bound of 1 ms [34], but close to the estimated ''speed limit'' [7].

We have shown that the dynamics of a reasonably complex model of protein folding can be embedded into a one-dimensional coordinate. This provides an explicit connection between theory and folding simulations and justifies the use of such coordinates in theory and the interpretation of experiments. The method should be useful in other high dimensional problems besides protein folding.

We thank W. A. Eaton, R. W. Pastor, A. Szabo, D. Thirumalai, and P. G. Wolynes for helpful comments on the manuscript. This research was supported by the Intramural Research Program of the NIH, NIDDK.

*Electronic address: best@helix.nih.gov
†Electronic address: hummer@helix.nih.gov

[1] N. D. Socci, J. N. Onuchic, and P. G. Wolynes, J. Chem. Phys. **104**, 5860 (1996).
[2] J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).
[3] A. Šali, E. Shaknovich, and M. Karplus, Nature (London) **369**, 248 (1994).
[4] C. J. Camacho and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **90**, 6369 (1993).
[5] K. A. Dill and H. S. Chan, Nat. Struct. Biol. **4**, 10 (1997).
[6] V. Muñoz and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **96**, 11 311 (1999).
[7] S. J. Hagen, J. Hofrichter, A. Szabo, and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **93**, 11 615 (1996).
[8] J. Sabelko, J. Ervin, and M. Gruebele, Proc. Natl. Acad. Sci. U.S.A. **96**, 6031 (1999).
[9] M. M. Garcia-Mira, M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz, and V. Muñoz, Science **298**, 2191 (2002).
[10] R. Zwanzig, Acc. Chem. Res. **23**, 148 (1990).
[11] W. Min and X. S. Xie, Phys. Rev. E **73**, 010902 (2006).
[12] H. Nymeyer, A. E. García, and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **95**, 5921 (1998).
[13] T. Wang, Y. Zhu, and F. Gai, J. Phys. Chem. B **108**, 3694 (2004).
[14] J. Karanicolas and C. L. Brooks, III, Protein Science **11**, 2351 (2002).
[15] R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U.S.A. **102**, 6732 (2005).
[16] D. J. Bicout and A. Szabo, J. Chem. Phys. **109**, 2325 (1998).
[17] G. Hummer, New J. Phys. **7**, 34 (2005).
[18] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[19] G. Hummer, J. Chem. Phys. **120**, 516 (2004).
[20] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[21] C. L. Brooks, III, A. Brünger, and M. Karplus, Chem. Phys. Lett. **105**, 495 (1984).
[22] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[23] W. F. van Gunsteren and H. J. C. Berendsen, Mol. Simul. **1**, 173 (1988).
[24] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, Comput. Phys. Commun. **91**, 43 (1995).
[25] H. A. Kramers, Physica (Amsterdam) **7**, 284 (1940).
[26] S. Chadrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[27] V. I. Mel'nikov and S. V. Meshkov, J. Chem. Phys. **85**, 1018 (1986).
[28] R. J. Loncharich, B. R. Brooks, and R. W. Pastor, Biopolymers **32**, 523 (1992).
[29] J. J. Portman, S. Takada, and P. G. Wolynes, J. Chem. Phys. **114**, 5082 (2001).
[30] J. E. Straub, M. Borkovec, and B. J. Berne, J. Chem. Phys. **84**, 1788 (1986).
[31] S. S. Plotkin and P. G. Wolynes, Phys. Rev. Lett. **80**, 5015 (1998).
[32] B. Zagrovic and V. S. Pande, J. Comput. Chem. **24**, 1432 (2003).
[33] D. K. Klimov and D. Thirumalai, Phys. Rev. Lett. **79**, 317 (1997).
[34] B. Schuler, E. A. Lipman, and W. A. Eaton, Nature (London) **419**, 743 (2002).