JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Are Peptides Good Two-State Folders?

Alexander M. Berezhkovskii,[†] Florentina Tofoleanu,[‡,§] and Nicolae-Viorel Buchete[*,‡,§]

[†]Mathematical and Statistical Computing Laboratory, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, United States

[‡]School of Physics and [§]Complex and Adaptive Systems Laboratory, University College Dublin, Belfield, Dublin 4, Ireland

ABSTRACT: The folding kinetics of proteins is frequently single-exponential, as basins of folded and unfolded conformations are well separated by a high barrier. However, for relatively short peptides, a two-state character of folding is rather the exception than the rule. In this work, we use a Zwanzig-type model of protein conformational dynamics to study the dependence of folding kinetics on the protein chain length, M. The analysis is focused on the gap in the eigenvalue spectrum of the rate matrix that describes the protein's conformational dynamics. When there is a large gap between the two smallest in magnitude nonzero eigenvalues, the corresponding relaxation times have qualitatively different physical interpretations. The longest of these two times characterizes the interbasin equilibration (i.e., folding), whereas the second time characterizes the intrabasin relaxation. We derive approximate analytical solutions for the two eigenvalues that show how they depend on M. From these solutions, we infer that there is a large gap between the two, and thus, the kinetics is essentially single-exponential when M is large enough such that 2^{M+1} is much larger than M^2.

INTRODUCTION

The kinetics of many long proteins is single-exponential. This implies that folding of such proteins is a two-state process that involves transitions between the folded and unfolded states of the protein, which are separated by a high barrier. However, this is not necessarily the case for short peptides. The goal of the present work is to study how the character of the folding kinetics (i.e., whether the kinetics is single-exponential or not) depends on the protein chain length. We analyzed this question within the framework of a simple model of the protein conformational dynamics that is similar to the model suggested by Zwanzig, Szabo, and Bagchi to study Levinthal's paradox.[1−4] Being simple, the model allows for an analytical solution, which is used to establish the relation between the character of the folding kinetics and the protein chain length. Our results are of particular interest to those researchers who study the folding of short peptides, either experimentally or computationally, with atomic-level detail or using simplified models,[5−24] as their results are frequently interpreted within the framework of simple two-state models. Many theoretical and experimental studies use small homoproteins such as dialanine (Ala2) or pentaalanine (Ala5, the smallest peptide that forms one full helical loop) as test systems.[25−37]

Our interest in this problem was initiated by a molecular dynamics (MD) study of the folding kinetics of pentaalanine in explicit solvent,[35,36] in which it was shown that the folding kinetics is better approximated by a four-state rather than by a two-state model. Figure 1 illustrates some of the most populated unfolded conformational states of Ala5, together with the all-helical folded state. As shown herein, a possible cause for the deviation from a single-exponential relaxation might be the relatively small number of unfolded conformational states available to small peptides. Note that a complex, non-two-state, length-dependent folding kinetics has been reported in experimental studies.[6−8,12,38−41]

To analyze the character of the folding kinetics, we derived approximate expressions for the first two nonzero eigenvalues of the rate matrix that describes the conformational dynamics of our model protein. These eigenvalues are important because a large gap between them is a fingerprint of the two-state character of the kinetics. With analytical expressions for the eigenvalues in hand, we analyzed how the gap depends on the chain length of the protein. We checked the accuracy of our analytical results by comparison with the exact eigenvalues of the rate matrix found numerically. The comparison showed that the two agree well if the peptide is not too short. For our model, we found that the kinetics is essentially single-exponential when the chain length M of the protein satisfies $2^{M+1} \gg M^2$. To be more precise, our model protein must contain nine or more residues to be a good two-state folder.

MODEL AND METHODS

Consider two-state protein folding described by the kinetic scheme

$$\mathrm{U} \underset{k_{\mathrm{U}}}{\overset{k_{\mathrm{F}}}{\rightleftharpoons}} \mathrm{F} \qquad (1)$$

where U and F denote the unfolded (denatured) and folded (native) states of the protein, respectively, and k_{F} and k_{U} are the corresponding folding and unfolding rate constants. To understand the mechanism of protein folding, one has to establish a relation between eq 1 and the underlying protein dynamics.

In our model, the protein is a homopolymer that contains M identical residues, each of which can be in either a folded (f) or an unfolded (u) state. To characterize the state of the residue, we introduce an "indicator", s, and assign $s = 0$ and 1 to the residue in states f and u, respectively. A protein conformation is completely characterized by a set $\{s\} = (s_1, s_2, ..., s_M)$, where s_i indicates the state of residue i, $i = 1, 2, ..., M$. The number of residues in state u

Received: April 22, 2011
Published: June 30, 2011

in a protein conformation $\{s\}$, $m(\{s\})$, is given by

$$m(\{s\}) = \sum_{i=1}^{M} s_i \tag{2}$$

Thus, $0 \le m \le M$. The conformation with no unfolded residues (i.e., $s_i = 0$ for all i, and hence, $m = 0$) represents the folded state of the protein. In all other conformations, $m > 0$, and the protein contains m residues in the u state and $M - m$ residues in the f state. The number of conformations with a fixed value m of unfolded residues is given by the binomial coefficient

$$\binom{M}{m} = \frac{M!}{m!(M-m)!}$$

and the total number of the conformations of our protein is

$$\sum_{m=0}^{M} \binom{M}{m} = 2^M$$

Let $P_m(t)$ be the probability of finding the protein in one of the conformations with m unfolded residues at time t, $\sum_{m=0}^{M} P_m(t) = 1$. We assume that the evolution of $P_m(t)$ is due to independent transitions between f and u states of individual residues. We take these transitions to be described by the kinetic scheme

$$\mathrm{u} \underset{k}{\overset{k}{\rightleftharpoons}} \mathrm{f} \tag{3}$$

where k is the rate constant, which is assumed to be the same for transitions in both directions. To introduce a collective behavior, responsible for the folding of our protein, we assume that the unfolding of the first and second residues occurs with the modified rate constants $\varepsilon_0 k$ and $\varepsilon_1 k$, respectively, with $\varepsilon_0, \varepsilon_1 \ll 1$. The precise choices of ε_0 and ε_1 are specified later. Figure 2A illustrates the connectivity network of the conformational dynamics of our peptide of length $M = 5$.

The evolution of the distribution function $P_m(t)$ is described by the master equation

$$\begin{aligned}
\dot{P}_0(t) &= -\varepsilon_0 M P_0(t) + P_1(t) \\
\dot{P}_1(t) &= \varepsilon_0 M P_0(t) - [1 + \varepsilon_1(M-1)]P_1(t) + 2P_2(t) \\
\dot{P}_2(t) &= \varepsilon_1(M-1)P_1(t) - MP_2(t) + 3P_3(t) \\
&\vdots \\
\dot{P}_m(t) &= (M - m + 1)P_{m-1}(t) - MP_m(t) \\
&\quad + (1 - \delta_{mM})(m+1)P_{m+1}(t), \text{ for } 3 \le m \le M
\end{aligned} \tag{4}$$

where the time t is expressed in units of $1/k$ and δ_{mn} is the Kronecker delta. This set of equations describes the relaxation of $P_m(t)$ to the equilibrium distribution P_m^{eq} given by

$$P_m^{eq} = \frac{[\varepsilon_0 + (1-\varepsilon_0)\delta_{m0}][\varepsilon_1 + (1-\varepsilon_1)\delta_{m1}]}{[1-(1-\varepsilon_1)\delta_{m0}][1 + \varepsilon_0(1-\varepsilon_1)M + \varepsilon_0\varepsilon_1(2^M - 1)]}\binom{M}{m} \tag{5}$$

One can check that P_m^{eq} satisfies eq 4 and the normalization condition, $\sum_{m=0}^{M} P_m^{\mathrm{eq}} = 1$.

To determine the parameters ε_0 and ε_1, we impose two requirements: (i) The probability of finding the protein in the native state ($m = 0$) is 0.5 (i.e., the calculations are done at the protein melting point), and (ii) the protein in conformations with one unfolded residue ($m = 1$) makes forward (i.e., $1 \rightarrow 2$) and backward (i.e., $1 \rightarrow 0$)

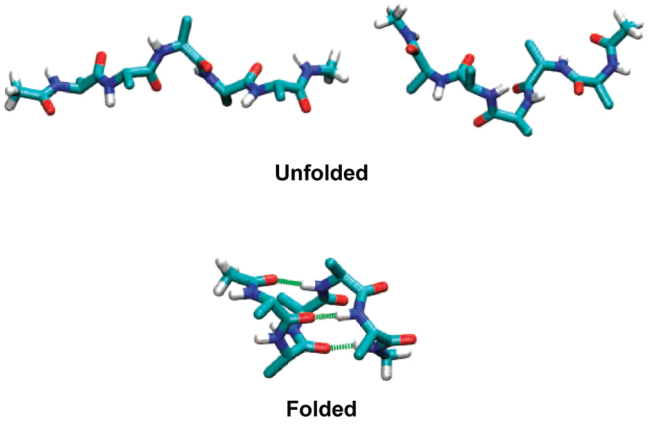

Figure 1. Folded and unfolded conformations of pentaalanine.

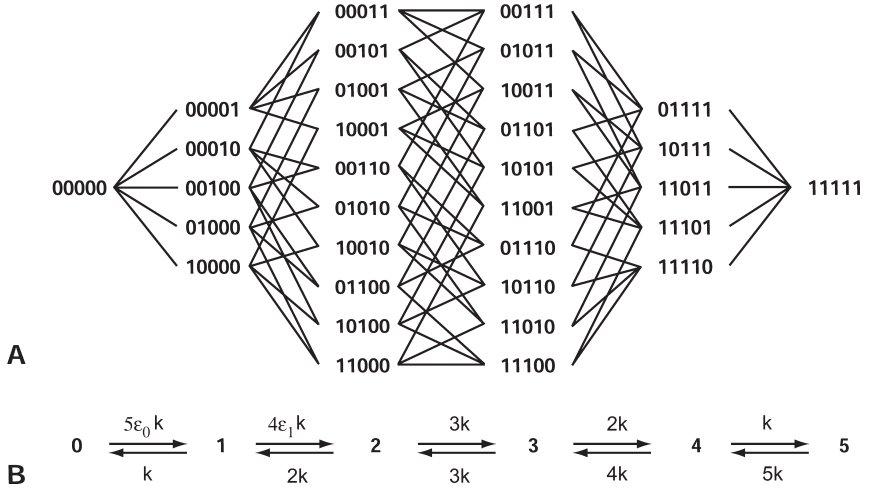

Figure 2. Schematic representation of conformational dynamics of our model peptide for length $M = 5$. (A) Connectivity network showing that only transitions between nearest neighbors are permitted. Conformations of the peptide are described using a binary notation for each residue, with 0 and 1 denoting folded and unfolded states of the residue. (B) Effective conformational dynamics along the discrete one-dimensional reaction coordinate, m.

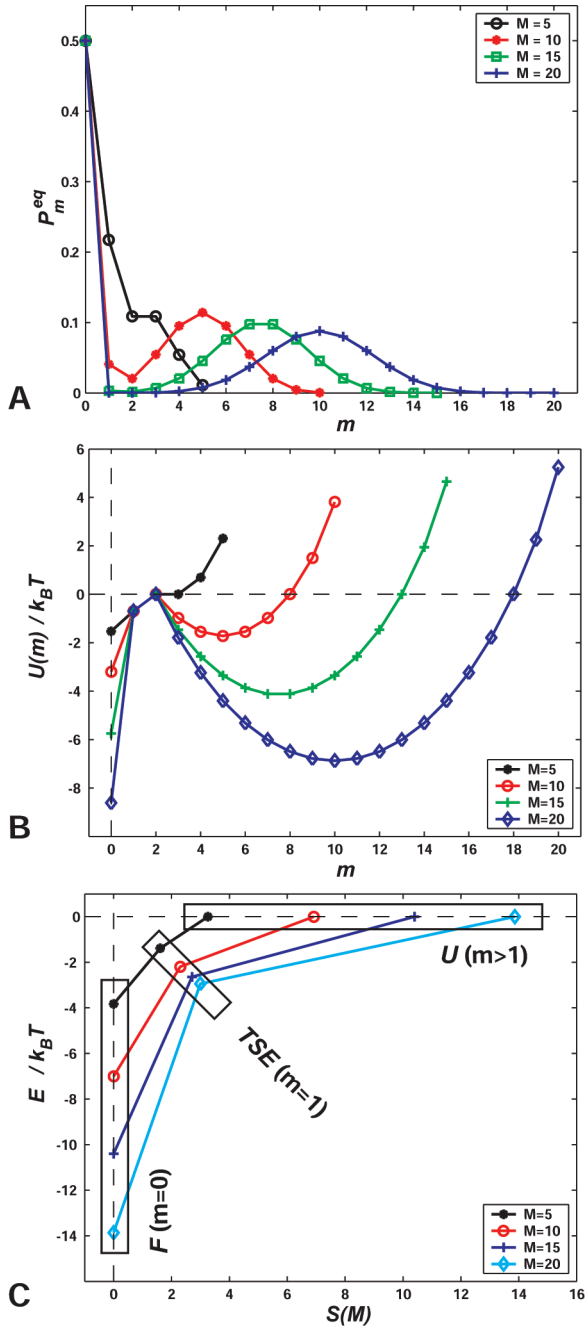

Figure 3. (A) Equilibrium distribution P_m^{eq} and (B) potential of mean force $U(m)$ as functions of the reaction coordinate m. (C) Folding funnel.

transitions with equal probability. In other words, conformations with $m = 1$ can be considered as transition states between the folded and unfolded conformations of the protein. These requirements lead to the following values of the parameters

$$\varepsilon_0 = \frac{M-1}{2^M + M^2 - 2M - 1}, \quad \varepsilon_1 = \frac{1}{M-1} \tag{6}$$

As a consequence, (i) the equilibrium distribution in eq 5 simplifies and takes the form

$$P_m^{eq} = \frac{[\varepsilon_0 + (1-\varepsilon_0)\delta_{m0}][\varepsilon_1 + (1-\varepsilon_1)\delta_{m1}]}{2[1-(1-\varepsilon_1)\delta_{m0}]}\binom{M}{m} \tag{7}$$

and (ii) the product $\varepsilon_1(M-1)$ in eq 4 becomes unity.

Figure 3 illustrates some properties of the proteins of different chain lengths M. Figure 3A shows the equilibrium distribution P_m^{eq}. In our model, the equilibrium probability of the folded state is always 0.5. Interestingly, the second peak (for the unfolded ensemble) appears only when $M > 5$. Using eq 7, one can check that P_m^{eq} has a minimum at $m = 2$ for all $M > 5$. As the size M of the peptide chain increases, the folded and unfolded basins of the peptide's chain become better separated, and the transition state conformations corresponding to $m = 1$ become less populated. In addition, the distribution of unfolded conformations becomes more and more Gaussian.

The value of m characterizes the "distance" of the conformation from the native state. We chose the number of unfolded residues, m, as a discrete one-dimensional reaction coordinate. The effective conformational dynamics along this coordinate is illustrated in Figure 2B. Note that m is similar to the fraction of native contacts, Q, which has been used as a reaction coordinate in numerous studies of protein folding.[42−45]

We use the equilibrium distribution, P_m^{eq}, to introduce the potential of mean force, $U(m)$, along the reaction coordinate

$$U(m) = -k_B T \ln \frac{P_m^{eq}}{P_2^{eq}} \tag{8}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and we choose $U(2)$ as the reference such that $U(2) = 0$ for all peptides. Figure 3B shows $U(m)$ for proteins of different lengths M. Note that (i) the potential of mean force has a barrier only when $M > 5$, and this barrier is always at $m = 2$; (ii) the transition state conformations with $m = 1$ have a lower energy that is independent of the chain length; and (iii) the barrier height increases quickly with M.

In our model, all unfolded (U) conformations with $m \geq 2$ have the same energy. This is a consequence of the identity of the forward and backward rate constants in eq 3. The energy of the conformations with $m = 1$, which form the transition state ensemble (TSE), is $k_BT \ln(1/\varepsilon_1) = k_BT \ln(M-1)$ lower than the energy of the unfolded conformations. Finally, the energy of the folded (F) conformation is $k_BT \ln[1/(\varepsilon_0\varepsilon_1)] = k_BT \ln(2^M + M^2 - 2M - 1)$ lower than that of the unfolded conformations. The dimensionless entropy of each of the three groups of conformations is defined as the natural logarithm of the number of conformations in the group

$$S_F = 0, \quad S_{TSE} = \ln M, \quad S_U = \ln(2^M - M - 1) \tag{9}$$

We use these relations to draw the folding funnels[42,43] for our model proteins of different chain lengths M, shown in Figure 3C, where we chose the energy of the unfolded conformations as zero. Note that the TSE conformations with $m = 1$ are in the funnel. As the chain length increases, the slope of the TSE-U part of the funnel decreases, whereas the slope of the TSE-F part increases.

RESULTS AND DISCUSSION

To analyze the character of the relaxation kinetics, one needs to know the eigenvalues of the evolution operator. In vector−matrix notation, the master equation in eq 4 takes the form

$$\dot{\mathbf{P}}(t) = \mathbf{K}\mathbf{P}(t) \tag{10}$$

Here, the evolution operator **K** is the $(M+1)\times(M+1)$ three-diagonal rate matrix

$$\mathbf{K}=\begin{pmatrix} -\alpha_0 & \beta_1 & 0 & \cdots & 0 & 0 & 0 \\ \alpha_0 & -\gamma_1 & \beta_2 & \cdots & 0 & 0 & 0 \\ 0 & \alpha_1 & -\gamma_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\gamma_{M-2} & \beta_{M-1} & 0 \\ 0 & 0 & 0 & \cdots & \alpha_{M-2} & -\gamma_{M-1} & \beta_M \\ 0 & 0 & 0 & \cdots & 0 & \alpha_{M-1} & -\beta_M \end{pmatrix} \tag{11}$$

with matrix elements given by

$$\begin{aligned} &\alpha_0=\varepsilon_0 M, \quad \alpha_1=1, \quad \alpha_m=M-m, \quad \text{for } m=2,3,\ldots,M-1 \\ &\beta_m=m, \quad \text{for } m=1,2\ldots,M \\ &\gamma_1=2, \quad \gamma_m=\alpha_m+\beta_m=M, \quad \text{for } m=2,\ldots,M-1 \end{aligned} \tag{12}$$

Eigenvalues, $-\lambda_i$, of the rate matrix and the corresponding eigenvectors, $\boldsymbol{\varphi}_i$, are solutions of the eigenvalue problem

$$\mathbf{K}\varphi_i=-\lambda_i\varphi_i, \quad i=1,2,\ldots,M+1 \tag{13}$$

The eigenvalue with the smallest magnitude is equal to zero, $\lambda_1=0$, because the master equation, eq 10, describes the relaxation of the initial distribution to equilibrium. We assume that the eigenvalues are ordered in increasing values of their magnitudes, $\lambda_1=0<\lambda_2\le\lambda_3\le\cdots\le\lambda_{M+1}$.

When the system consists of two basins separated by a high barrier, intrabasin relaxation occurs much faster than interbasin exchange. As a consequence, the exchange is a memory-less Markov process, and hence, equilibration of basin's populations is single-exponential. The eigenvalues of the evolution operator of such a system have a simple physical interpretation. The smallest nonzero eigenvalue, $-\lambda_2$, describes interbasin equilibration of the populations, whereas the other nonzero eigenvalues describe intrabasin relaxation to local equilibria in the two basins. The fact that the former process is much slower than the latter leads to the inequality $\lambda_2\ll\lambda_3$, which implies that there is a large gap in the eigenvalue spectrum of the evolution operator.

For our simple model, one can find approximate solutions for λ_2 and λ_3 (see Appendix A)

$$\lambda_2\approx\varepsilon_0 M=\frac{M(M-1)}{2^M+M^2-2M-1}, \quad \lambda_3\approx 2 \tag{14}$$

This allows us to establish the condition of applicability of the two-state description of folding

$$\lambda_3\gg\lambda_2\Rightarrow 2^{M+1}\gg M^2 \tag{15}$$

which is the main result of the present study. We show below that the relaxation is single-exponential to a good approximation for $M\ge 9$.

In our model, $k_F=k_U=\lambda_2/2$, and according to eq 14, at large M, we have $\ln\tau_F\propto M$, where $\tau_F=1/k_F$ is the folding time. Note that linear scaling of the logarithm of the folding time with the protein size has been reported in the literature.[46,47] A weaker scaling with the chain length, $\ln\tau_F\propto M^\gamma$, where $\gamma<1$, has also been discussed.[47−52]

Finally, we compare the values of λ_2 and λ_3 given in eq 14 to the eigenvalues found numerically by diagonalization of the rate matrix (eq 11). The ratios of the eigenvalues for proteins of different lengths are presented in Table 1. One can see that eq 14 provides good estimations for λ_2 and λ_3 when M is large enough. As expected, the agreement gets better as M increases. In Table 2, the ratios λ_3/λ_2 obtained numerically are compared with $2^{M+1}/M^2$, which is the large-M estimation of the ratio that follows from eq 14. Comparison shows that the simple expression $2^{M+1}/M^2$ appears to be a good approximation for the ratio of the eigenvalues.

Table 1. Ratios of the Approximate Results in eq 14, $\lambda_{2,3}^{app}$, to Their Counterparts Obtained by Numerical Diagonalization of the Rate Matrix in eq 11, $\lambda_{2,3}^{diag}$

	M					
	5	10	20	30	40	50
$\lambda_2^{app}/\lambda_2^{diag}$	0.86	1.13	1.07	1.04	1.03	1.01
$\lambda_3^{app}/\lambda_3^{diag}$	1.02	1.24	1.09	1.04	1.03	1.02

One can interpret λ_2^{-1} and λ_3^{-1} as slow and fast relaxation times of the system, $\tau_{slow}=1/\lambda_2$, $\tau_{fast}=1/\lambda_3$. Figure 4 shows the ratio $(\tau_{slow}-\tau_{fast})/\tau_{fast}$, which characterizes the gap between the relaxation times, as a function of the chain length. One can see that the gap quickly increases with M and exceeds 10 for $M\ge 9$. Written in terms of λ_2 and λ_3, the ratio takes the form $(\tau_{slow}-\tau_{fast})/\tau_{fast}=\lambda_3/\lambda_2-1$. When M is large, $\lambda_3/\lambda_2\approx 2^{M+1}/M^2$, so that $(\tau_{slow}-\tau_{fast})/\tau_{fast}\approx 2^{M+1}/M^2$. Figure 4 shows that $2^{M+1}/M^2$ is close to the exact values of the ratio obtained numerically even at not-so-large values of M.

In summary, a simple model of the conformational dynamics has been used to analyze how the character of the folding kinetics depends on the chain length of the protein. We found that, to a good approximation, a protein can be considered as a two-state folder when the number of its residues is nine or larger. In view of this estimation, it is not surprising that the folding kinetics of pentaalanine deviates from single-exponential behavior.[35,36]

APPENDIX A

When M is large, the potential of mean force $U(m)$ has a well-pronounced double-well structure with a high barrier that separates folded and unfolded conformations (Figure 3B). We assume that equilibration in the basin of unfolded conformations, $m=2,3,\ldots,M$, is a fast process, $\lambda_4,\lambda_5,\ldots,\lambda_{M+1}\gg 1$, so that this basin is always in local equilibrium and

$$P_m(t)=\frac{P_m^{eq}}{P_2^{eq}}P_2(t)=\frac{2(M-2)!}{m!(M-m)!}P_2(t), \quad m=2,3,\ldots,M \tag{16}$$

The probability of finding the protein in the unfolded basin at time t, $P_U(t)$, is

$$P_U(t)=\sum_{m=2}^{M}P_m(t)=\frac{2(2^M-M-1)}{M(M-1)}P_2(t) \tag{17}$$

Using this expression and the fact that $P_U(t)=1-P_0(t)-P_1(t)$, we can write $P_2(t)$ in terms of $P_0(t)$ and $P_1(t)$

$$\begin{aligned} P_2(t)&=\frac{M(M-1)}{2(2^M-M-1)}P_U(t) \\ &=\frac{M(M-1)}{2(2^M-M-1)}[1-P_0(t)-P_1(t)] \end{aligned} \tag{18}$$

dx.doi.org/10.1021/ct200281d |J. Chem. Theory Comput. 2011, 7, 2370–2375

Table 2. Ratio of the First Two Nonzero Eigenvalues Obtained by Numerical Diagonalization of the Rate Matrix in eq 11, $\lambda_3^{diag}/\lambda_2^{diag}$, and the Large-$M$ Estimation of This Ratio, $2^{M+1}/M^2$, as Functions of the Chain Length

	M					
	5	10	20	30	40	50
$\lambda_3^{diag}/\lambda_2^{diag}$	3.89	22.21	5.43×10^3	2.46×10^6	1.41×10^9	9.08×10^{11}
$2^{M+1}/M^2$	2.56	20.48	5.24×10^3	2.39×10^6	1.37×10^9	9.01×10^{11}

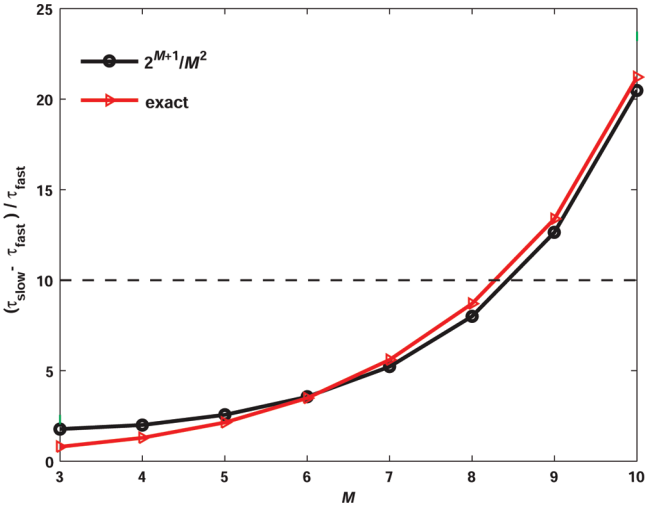

Figure 4. Ratio of the relaxation times $(\tau_{slow} - \tau_{fast})/\tau_{fast}$ and its large-M estimation, $2^{M+1}/M^2$, as functions of the chain length. The exact values (triangles) were obtained by diagonalization of the rate matrix.

Substituting this expression into the second equation of the set in eq 4, we find that the first two equations of this set form closed evolution equations for the probabilities $P_0(t)$ and $P_1(t)$

$$\dot{P}_0(t) = -\varepsilon_0 M P_0(t) + P_1(t)$$
$$\dot{P}_1(t) = \varepsilon_0 M P_0(t) - 2P_1(t) + \frac{M(M-1)}{2^M - M - 1}[1 - P_0(t) - P_1(t)] \quad (19)$$

where we have used the relation $\varepsilon_1(M-1) = 1$ (eq 6).

Equations 19 describe relaxation of $P_0(t)$ and $P_1(t)$ to their equilibrium values (eq 7)

$$P_0^{eq} = 1/2, \quad P_1^{eq} = \varepsilon_0 M P_0^{eq} = \frac{M(M-1)}{2(2^M + M^2 - 2M - 1)} \quad (20)$$

Denoting the deviation of $P_i(t)$ from P_i^{eq} by $\Delta P_i(t)$, $\Delta P_i(t) = P_i(t) - P_i^{eq}(t)$, $i = 0, 1$, and using eq 19, we find that the deviations satisfy

$$\Delta\dot{P}_0(t) = -\varepsilon_0 M \Delta P_0(t) + \Delta P_1(t)$$
$$\Delta\dot{P}_1(t) = \left[\varepsilon_0 M - \frac{M(M-1)}{2^M - M - 1}\right]\Delta P_0(t) - \left[2 + \frac{M(M-1)}{2^M - M - 1}\right]\Delta P_1(t) \quad (21)$$

The second equation can be simplified at large M, because the first term on the right-hand side can be neglected as the factor in front of $\Delta P_0(t)$ is proportional to $\varepsilon_0^2 \ll 1$, whereas the factor in front of $\Delta P_1(t)$ is close to 2. As a result, the second equation takes the form

$$\Delta\dot{P}_1(t) = -2\Delta P_1(t) \quad (22)$$

From eqs 21 and 22, one can see that the relaxation to equilibrium is biexponential and that λ_2 and λ_3 are given by the expressions in eq 14.

AUTHOR INFORMATION

Corresponding Author
*E-mail: buchete@ucd.ie.

ACKNOWLEDGMENT

We thank Gerhard Hummer and Attila Szabo for many helpful and stimulating discussions. This study was supported by the Intramural Research Program of the National Institutes of Health (NIH), Center for Information Technology, and it used the computational resources of the Irish Centre for High-End Computing (ICHEC). F.T. and N.V.B gratefully acknowledge financial support from the Irish Research Council for Science, Engineering & Technology (IRCSET).

REFERENCES

(1) Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's Paradox. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 20.
(2) Zwanzig, R. Simple Model of Protein Folding Kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9801.
(3) Zwanzig, R. Two-State Models of Protein Folding Kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 148.
(4) Buchete, N. V.; Straub, J. E. Mean First-Passage Time Calculations for the Coil-to-Helix Transition: The Active Helix Ising Model. *J. Phys. Chem. B* **2001**, *105*, 6684.
(5) Go, M.; Go, N.; Scheraga, H. A. Molecular Theory of the Helix−Coil Transition in Polyamino Acids. 3. Evaluation and Analysis of S and Sigma for Polyglycine and Poly-L-Alanine in Water. *J. Chem. Phys.* **1971**, *54*, 4489.
(6) Huang, C. Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai, F. Temperature-Dependent Helix−Coil Transition of an Alanine Based Peptide. *J. Am. Chem. Soc.* **2001**, *123*, 9235.
(7) Wang, T.; Du, D. G.; Gai, F. Helix−Coil Kinetics of Two 14-Residue Peptides. *Chem. Phys. Lett.* **2003**, *370*, 842.
(8) Wang, T.; Zhu, Y.; Getahun, Z.; Du, D.; Huang, C.-Y.; DeGrado, W. F.; Gai, F. Length Dependent Helix−Coil Transition Kinetics of Nine Alanine-Based Peptides. *J. Phys. Chem. B* **2004**, *108*, 15301.
(9) Buchete, N. V.; Straub, J. E.; Thirumalai, D. Orientation-Dependent Coarse-Grained Potentials Derived by Statistical Analysis of Molecular Structural Databases. *Polymer* **2004**, *45*, 597.
(10) Buchete, N. V.; Straub, J. E.; Thirumalai, D. Continuous Anisotropic Representation of Coarse-Grained Potentials for Proteins by Spherical Harmonics Synthesis. *J. Mol. Graph.* **2004**, *22*, 441.
(11) Van Giessen, A. E.; Straub, J. E. Coarse-Grained Model of Coil-to-Helix Kinetics Demonstrates the Importance of Multiple Nucleation Sites in Helix Folding. *J. Chem. Theory Comput.* **2006**, *2*, 674.
(12) Mukherjee, S.; Chowdhury, P.; Bunagan, M. R.; Gai, F. Folding Kinetics of a Naturally Occurring Helical Peptide: Implication of the Folding Speed Limit of Helical Proteins. *J. Phys. Chem. B* **2008**, *112*, 9146.
(13) Buchete, N. V.; Straub, J. E.; Thirumalai, D. Dissecting Contact Potentials for Proteins: Relative Contributions of Individual Amino Acids. *Proteins* **2008**, *70*, 119.

(14) Rosta, E.; Buchete, N. V.; Hummer, G. Thermostat Artifacts in Replica Exchange Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 1393.
(15) Freddolino, P. L.; Harrison, C. B.; Liu, Y. X.; Schulten, K. Challenges in Protein-Folding Simulations. *Nat. Phys.* **2010**, *6*, 751.
(16) Li, M. S.; Klimov, D. K.; Thirumalai, D. Finite Size Effects on Calorimetric Cooperativity of Two-State Proteins. *Physica A* **2005**, *350*, 38.
(17) Berezhkovskii, A.; Szabo, A. Perturbation Theory of Φ-Value Analysis of Two-State Protein Folding: Relation between p_{fold} and Φ Values. *J. Chem. Phys.* **2006**, *125*, 104902.
(18) Muñoz, V.; Eaton, W. A. A Simple Model for Calculating the Kinetics of Protein Folding from Three-Dimensional Structures. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11311.
(19) Muñoz, V. Thermodynamics and Kinetics of Downhill Protein Folding Investigated with a Simple Statistical Mechanical Model. *Int. J. Quantum Chem.* **2002**, *90*, 1522.
(20) Bruscolini, P.; Pelizzola, A. Exact Solution of the Muñoz-Eaton Model for Protein Folding. *Phys. Rev. Lett.* **2002**, *88*, 258101.
(21) Ivankov, D. N.; Garbuzynskiy, S. O.; Alm, E.; Plaxco, K. W.; Baker, D.; Finkelstein, A. V. Contact Order Revisited: Influence of Protein Size on the Folding Rate. *Protein Sci.* **2003**, *12*, 2057.
(22) Weikl, T. R.; Palassini, M.; Dill, K. A. Cooperativity in Two-State Protein Folding Kinetics. *Protein Sci.* **2004**, *13*, 822.
(23) Ferguson, A.; Liu, Z.; Chan, H. S. Desolvation Barrier Effects Are a Likely Contributor to the Remarkable Diversity in the Folding Rates of Small Proteins. *J. Mol. Biol.* **2009**, *389*, 619.
(24) Murza, A.; Kubelka, J. Beyond the Nearest-Neighbor Zimm−Bragg Model for Helix−Coil Transition in Peptides. *Biopolymers* **2009**, *91*, 120.
(25) Brooks, C. L.; Case, D. A. Simulations of Peptide Conformational Dynamics and Thermodynamics. *Chem. Rev.* **1993**, *93*, 2487.
(26) Brooks, C. L. Helix−Coil Kinetics: Folding Time Scales for Helical Peptides from a Sequential Kinetic Model. *J. Phys. Chem.* **1996**, *100*, 2546.
(27) Hummer, G.; Garcia, A. E.; Garde, S. Helix Nucleation Kinetics from Molecular Simulations in Explicit Solvent. *Proteins* **2001**, *42*, 77.
(28) Margulis, C. J.; Stern, H. A.; Berne, B. J. Helix Unfolding and Intramolecular Hydrogen Bond Dynamics in Small α-Helices in Explicit Solvent. *J. Phys. Chem. B* **2002**, *106*, 10748.
(29) Hummer, G.; Kevrekidis, I. G. Coarse Molecular Dynamics of a Peptide Fragment: Free Energy, Kinetics, and Long-Time Dynamics Computations. *J. Chem. Phys.* **2003**, *118*, 10762.
(30) Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. *J. Am. Chem. Soc.* **2007**, *129*, 1179.
(31) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov Models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* **2007**, *126*, 155101.
(32) Wickstrom, L.; Okur, A.; Simmerling, C. Evaluating the Performance of the ff99SB Force Field Based on NMR Scalar Coupling Data. *Biophys. J.* **2009**, *97*, 853.
(33) Ruzhytska, S.; Jacobi, M. N.; Jensen, C. H.; Nerukh, D. Identification of Metastable States in Peptide's Dynamics. *J. Chem. Phys.* **2010**, *133*, 164102.
(34) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* **2010**, *78*, 1950.
(35) Buchete, N. V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *J. Phys. Chem. B* **2008**, *112*, 6057.
(36) Buchete, N. V.; Hummer, G. Peptide Folding Kinetics from Replica Exchange Molecular Dynamics. *Phys. Rev. E* **2008**, *77*, 030902.
(37) Buchner, G. S.; Murphy, R. D.; Buchete, N. V.; Kubelka, J. Dynamics of Protein Folding: Probing the Kinetic Network of Folding−Unfolding Transitions with Experiment and Theory. *Biochim. Biophys. Acta* **2011**, *1814*, 1001.
(38) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. Laser Temperature Jump Study of the Helix−Coil Kinetics of an Alanine Peptide Interpreted with a 'Kinetic Zipper' Model. *Biochemistry* **1997**, *36*, 9200.
(39) Yang, W. Y.; Gruebele, M. Folding at the Speed Limit. *Nature* **2003**, *423*, 193.
(40) Kubelka, J.; Hofrichter, J.; Eaton, W. A. The Protein Folding 'Speed Limit'. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76.
(41) Amunson, K. E.; Ackels, L.; Kubelka, J. Site-Specific Unfolding Thermodynamics of a Helix-Turn-Helix Protein. *J. Am. Chem. Soc.* **2008**, *130*, 8146.
(42) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. Navigating the Folding Routes. *Science* **1995**, *267*, 1619.
(43) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels. *J. Chem. Phys.* **1996**, *104*, 5860.
(44) Best, R. B.; Hummer, G. Reaction Coordinates and Rates from Transition Paths. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6732.
(45) Best, R. B.; Hummer, G. Coordinate-Dependent Diffusion in Protein Folding. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1088.
(46) Plotkin, S. S.; Wang, J.; Wolynes, P. G. Statistical Mechanics of a Correlated Energy Landscape Model for Protein Folding Funnels. *J. Chem. Phys.* **1997**, *106*, 2932.
(47) Gutin, A. M.; Abkevich, V. I.; Shakhnovich, E. I. Chain Length Scaling of Protein Folding Time. *Phys. Rev. Lett.* **1996**, *77*, 5433.
(48) Thirumalai, D. From Minimal Models to Real Proteins: Time Scales for Protein Folding Kinetics. *J. Phys. I* **1995**, *5*, 1457.
(49) Finkelstein, A. V.; Badretdinov, A. Y. Rate of Protein Folding near the Point of Thermodynamic Equilibrium between the Coil and the Most Stable Chain Fold. *Fold. Des.* **1997**, *2*, 115.
(50) Finkelstein, A. V.; Badretdinov, A. Y. Influence of Chain Knotting on the Rate of Folding. *Fold. Des.* **1998**, *3*, 67.
(51) Wolynes, P. G. Folding Funnels and Energy Landscapes of Larger Proteins within the Capillarity Approximation. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 6170.
(52) Naganathan, A. N.; Muñoz, V. Scaling of Folding Times with Protein Size. *J. Am. Chem. Soc.* **2005**, *127*, 480.