Unfolding rates for the diffusion-collision model

Chris Beck and Xavier Siemens
Molecular Modeling Laboratory, Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155
(Received 13 June 2001; published 24 October 2001)

In the diffusion-collision model, the unfolding rates are given by the likelihood of secondary structural cluster dissociation. In this work, we introduce an unfolding rate calculation for proteins whose secondary structural elements are α helices, modeled from thermal escape over a barrier that arises from the free energy in buried hydrophobic residues. Our results are in good agreement with currently accepted values for the attempt rate.

DOI: 10.1103/PhysRevE.64.052902 PACS number(s): 87.14.Ee, 87.15.Aa, 82.20.Db

I. INTRODUCTION

In the diffusion-collision model of protein folding [1] the protein is modeled using a collection of spheres connected by flexible strings. The spheres represent the secondary structural elements such as α helices or β sheets (or clusters of these secondary structures) called microdomains, that constitute the protein.

The folding process from a completely unfolded protein to the the final native state is accomplished via diffusion through the solvent, collision, and finally coalescence of the microdomains. The state of the protein is defined by the number of pairings between the microdomains that are present at a given time t. The rate equations for transitions between these states can be written as

$$\frac{d\mathbf{P}(t)}{dt}=\hat{K}\mathbf{P}(t), \tag{1}$$

where $\mathbf{P}(t)$ is the vector of states and $\hat{K}$ is a matrix containing the transition rates between the different states. A protein having, say, q microdomains would involve $p=q(q-1)/2$ pairings, 2^p states $P_i(t)$, and a $2^p\times 2^p$ rate matrix $\hat{K}$.

In general, the calculation of the elements of the rate matrix $\hat{K}$ is somewhat involved. The forward rates are the rates of structural coalescence. In the diffusion-collision model the forward rates are calculated assuming the microdomains diffuse through a solvent environment, the space of which is limited by the length of the intervening strings and the van der Waals radii of the microdomains. These microdomains are assumed to be nascently formed, and their degree of formation is given by a helix-coil transition theory calculation [2] (as in AGADIR [3,4]) in the case of α helices, or via a combination of theory [5] and experiment [6] in the case of β sheets. As the microdomains undergo diffusion, they occasionally collide. When this happens the microdomains coalesce with a probability γ, being held together by hydrophobic interactions in the case of α helices, or a combination of hydrophobic and hydrogen bond interactions in the case of β sheets. The coalescence probability γ is given by the likelihood that the microdomain is in α helical or β sheet form, the percentage of hydrophobic area, and the likelihood of proper geometrical orientation upon collision.

The forward folding times between any two given states in the mean first passage time approximation [7] are given by

$$\tau_f=\frac{l^2}{D}+\frac{LV(1-\gamma)}{\gamma DA}, \tag{2}$$

where V is the diffusion volume available to the microdomain pair, A is the target area for collisions, D is the relative diffusion coefficient, γ is the probability of coalescence upon collision, and l and L are geometrical parameters calculated for diffusion in a spherical space. The inverse of the first passage time scales τ_f are the forward folding rates k_f that are used in the rate matrix $\hat{K}$.

The pairs can also dissociate. In typical diffusion-collision model calculations, the form of the unfolding times τ_b used for two microdomains A and B comes from the Van't Hoff-Arrhenius law given by

$$\tau_b=\nu^{-1}e^{\Delta G_{AB}/k_BT}, \tag{3}$$

where ΔG_{AB} is the free energy difference between paired and unpaired states, k_B is Boltzmann's constant, T is the temperature, and ν is an attempt rate. In the case of α helices the dominant contribution to the free energy comes from the buried hydrophobic area, and therefore,

$$\Delta G_{AB}=fA_{AB}, \tag{4}$$

where f is the free energy change per unit buried hydrophobic area in the pairing [8] and A_{AB} is the buried area [9]. The unfolding rates k_b are given by the inverse of the unfolding times τ_b.

The diffusion-collision model has been successful in describing the overall folding kinetics of several proteins [10–12]. In each of these studies a single value of the parameter ν was used for every unfolding transition. This value was adjusted to obtain the desired result, namely, to ensure that the protein would fold to its native state. This procedure is justified because the equilibrium (or native) occupation probabilities are known; in fact, for sufficiently simple systems the folding and unfolding rates can be determined from these probabilities [13]. The typical values used lie between 1 and 1000 ns^{-1}, which yields unfolding rates consistent with observed rates of bimolecular dissociation [14].

In cases where the final occupation probabilities are unknown, for instance in the studies of protein misfolding and non-native kinetic intermediates [15] such methods are clearly not possible. Indeed, even a detailed description of the intermediate folding kinetics of a protein whose final

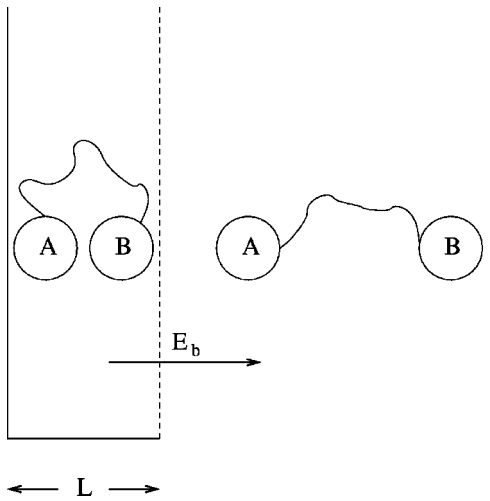

FIG. 1. Potential for the two microdomains *A* and *B*. The potential is infinite on the left because of the hard-core repulsion of the van der Walls contact between the microdomains. The barrier on the right can be crossed by microdomain pairs with energies larger than $E_b = fA_{AB}$, the free energy difference between paired and unpaired states with a buried hydrophobic area A_{AB}. The width of the well L, is taken to be the diameter of a water molecule.

state *is* known requires a more accurate and foundational determination of ν, as was pointed out by Burton *et al.* [10].

In this work we compute unfolding rates that, in the context of the diffusion-collision model, can be used for any given unfolding transition in the study of proteins whose secondary structural elements are α helices. From the rates we find the values of the parameter ν. This makes the diffusion-collision model more predictive and enables it to be used in situations where the occupation probabilities are unknown.

II. CALCULATION OF THE UNFOLDING RATES

We model the dissociation of microdomains as a thermal escape event over a barrier. Consider the pair of microdomains (which could be α helices or clusters of α helices) *A* and *B* connected by a string, diffusing in the potential well depicted in Fig. 1. The left boundary is infinite because of the hard-core repulsion of the van der Waals contact between the pair. Pairs with energies larger than $E_b = fA_{AB}$, the free energy difference between paired and unpaired states, can escape from the right boundary of the well. The well width L is set to the diameter of a water molecule. A separation larger than L exposes the buried hydrophobic area of the pair to the solvent, the free energy savings is lost, and the pair separates, resulting in an escape from the potential well.

The binding energies E_b of microdomain pairs in proteins are typically much larger than the thermal energy

$$E_b \gg k_B T. \tag{5}$$

This means that the time to escape from the well is much larger than any other time scale involved in the problem, in particular larger than the thermalization (or velocity autocorrelation) time and larger than the time it takes for the pair to diffuse in the well. Consequently, at any one time, the spatial distribution inside the well of an ensemble of pairs will be homogeneous

$$\rho(x,t) \propto 1/L \tag{6}$$

and the flux incident on the barriers will be thermal. We will use these two facts to calculate the rate at which the pairs dissociate.

The flux at the boundary on the right (at $x = L$) depends on the density of pairs at that boundary and the probability that their energy is high enough to thermally escape over the boundary. The differential element of flux at the boundary L of pairs with relative velocity between v and $v + dv$ is given by

$$dJ^{out}(L,t) = v\rho(L,t)dN(v), \tag{7}$$

where $\rho(L,t)$ is the number density of pairs at the boundary at a time t,

$$dN(v) = \left(\frac{\mu}{2\pi k_B T}\right)^{1/2} e^{-\mu v^2/2k_B T} dv \tag{8}$$

is the fraction of pairs with relative velocities between v and $v + dv$, and μ is the reduced mass given by

$$\mu = \frac{m_A m_B}{m_A + m_B}, \tag{9}$$

where m_A and m_B are the masses of the two microdomains.

In order to find the total flux through the outer boundary at L we must integrate over all velocities larger than $+\sqrt{E_b/2m}$ since the potential barrier can be crossed by pairs with energies higher than E_b, and pairs with relative velocities higher than that can contribute to the flux leaving the well. This yields a flux out of the well

$$J^{out}(L,t) = \rho(L,t)\left(\frac{k_b T}{2\pi\mu}\right)^{1/2} e^{-E_b/k_B T}. \tag{10}$$

If the number of pairs inside the well at some time t is $n(t)$ then, because of Eq. (5), the number density must be $\rho(x,t) = n(t)/L$ everywhere and

$$J^{out}(L,t) = \frac{n(t)}{L}\left(\frac{k_b T}{2\pi\mu}\right)^{1/2} e^{-E_b/k_B T}. \tag{11}$$

This means that the dissociation rate constant for a pair of microdomains with reduced mass μ and buried hydrophobic area $A_{AB} = E_b/f$ at a temperature T is

$$k_b = \frac{1}{L}\left(\frac{k_B T}{2\pi\mu}\right)^{1/2} e^{-E_b/k_B T}. \tag{12}$$

The terms preceding the exponential correspond to our prediction for the Van't Hoff-Arrhenius attempt rate ν in Eq. (3). As an example, the attempt rate found for a coalesced pair of 16-residue Regan-Degrado [16] helices with a combined hydrophobic area loss of 600 Å^2 is 64×10^9 s^{-1}.

It is interesting to note that a result similar to Eq. (12) would have been obtained by assuming the attempt rate to be the inverse of the thermal well-crossing time, namely taking

$$\nu \sim \frac{1}{L}\sqrt{\frac{k_B T}{\mu}} \tag{13}$$

in Eq. (3). This is, in fact, not the origin of the prefactors in Eq. (12). They arise as a consequence of Eq. (5): The factor of $1/L$ comes from the homogeneity of the spatial distribution (6) and the factor of $\sqrt{k_B T/2\pi\mu}$ from the integration of the thermal velocity distribution (8).

It is possible that dissociation events within a protein also include a relative rolling and/or sliding motion of the microdomains. In this case the calculation above can be performed with a few minor differences that take into account the extra degrees of freedom. The relative velocity distribution of the microdomains is still the one-dimensional Maxwell-Boltzmann distribution because motion parallel to the surface through which the probability is flowing does not contribute to escape from the well. The probability in the bound region is homogeneously distributed in a two- or three-dimensional volume in Eq. (10), and flows out of that volume through a one- or two-dimensional area. This calculation yields the result

$$k_b = \frac{d}{L}\left(\frac{k_b T}{2\pi\mu}\right)^{1/2} e^{-E_b/k_B T}, \tag{14}$$

where we set $d=2$ if we include either the rolling *or* sliding degrees of freedom, and $d=3$ if both of them are included. Due to the steric clashing of the side chains it seems rather unlikely that dissociation would include a sliding motion along the axes of the microdomains. It may be relevant, however, in the context of molten globules.

This approach succeeds in removing the free parameter ν from the diffusion-collision model. Moreover, our results for the one-, two-, and three-dimensional unfolding rates have a $\sqrt{T/\mu}$ dependence that could be used to distinguish between this and other proposals for the mechanism of microdomain pair dissociation.

III. CONCLUDING REMARKS

We have presented a calculation for the dissociation rate of a microdomain pair using a simple potential barrier over which pairs having energies above the free energy of the hydrophobic docking can escape. Since we have not accounted for the energy in hydrogen bonds this result is relevant for the dissociation of α-helix pairs or helix cluster pairs only and not for the dissociation of β-sheet pairs. We have found the unfolding rates arising from thermal fluctuations out of this potential well to be in good agreement with currently accepted values of the attempt rate ν.

The motivation of this work was to eliminate the free parameter ν from the diffusion-collision model. In previous applications of the diffusion-collision model (see for example, [10–12]) the folding kinetics from a denatured or random coil state to the final native state were followed. In such a case, it is reasonable to set the parameter ν such that the native state achieves most of the probability, because we know that the final state is attained at the end of the folding process. However, the removal of this parameter is important when considering folding processes that do not involve the native state. For example, in studying intermediate processes or protein misfolding [15], where the occupation probabilities may be completely unknown, such reasonable estimates of ν are not available. In these cases, elimination of ν as a free parameter is crucial.

The results presented here also predict a $\nu \propto \sqrt{T/\mu}$ dependence in all cases that can be distinguished experimentally from other possibilities such as $\nu \propto T$ [14]. Another difference is the dependence of the unfolding rates on the states, not only through the hydrophobic area, but also through the reduced mass μ of the microdomains undergoing dissociation. This is markedly different from typical diffusion-collision model calculations where the attempt rate ν is assumed to be the same for all dissociation events within the protein.

ACKNOWLEDGMENTS

We would like to thank Ken Olum, David Weaver, and Larry Ford for many useful discussions.

[1] M. Karplus and D.L. Weaver, Nature (London) **260**, 404 (1976); Biopolymers **18**, 1421 (1979).
[2] B.H. Zimm and J.K. Bragg, J. Chem. Phys. **31**, 526 (1959).
[3] V. Munoz and L. Serrano, Nature (London) **1**, 399 (1994); J. Mol. Biol. **245**, 275 (1994); **245**, 297 (1994); Biopolymers **41**, 495 (1997).
[4] E. Lacroix, A.R. Viguera, and L. Serrano, J. Mol. Biol. **284**, 173 (1998).
[5] V. Munoz, E.R. Henry, J. Hofrichter, and W.A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **95**, 5872 (1998).
[6] A.R. Dinner, T. Lazaridis, and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **96**, 9068 (1999).
[7] G.H. Weiss, Adv. Chem. Phys. **13**, 1 (1967); D.L. Weaver, Phys. Rev. B **20**, 2558 (1979); A. Szabo, K. Schulten, and Z. Schulten, J. Chem. Phys. **72**, 4350 (1980).
[8] C. Chothia, Nature (London) **248**, 338 (1974).
[9] B. Lee and F.M. Richards, J. Mol. Biol. **55**, 379 (1971).
[10] R.E. Burton, J.K. Meyers, and T.G. Oas, Biochemistry **37**, 5337 (1998).
[11] J.K. Meyers and T.G. Oas, J. Mol. Biol. **289**, 205 (1999).
[12] R.V. Pappu and D.L. Weaver, Protein Sci. **7**, 480 (1998).
[13] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[14] A. Fersht, *Structure and Mechanism in Protein Science* (Freeman, New York, 1999).
[15] C. Beck, X. Siemens, and D.L. Weaver, Biophy. J. (to be published).
[16] L. Regan and W.F. Degrado, Science **241**, 976 (1988).