

# The Origin of Nonmonotonic Complex Behavior and the Effects of Nonnative Interactions on the Diffusive Properties of Protein Folding

Ronaldo J. Oliveira,<sup>†△</sup> Paul C. Whitford,<sup>‡§\*\*△\*</sup> Jorge Chahine,<sup>†</sup> Jin Wang,<sup>¶||</sup> José N. Onuchic,<sup>\*\*</sup> and Vitor B. P. Leite<sup>†\*</sup>

<sup>†</sup>Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, Brazil;

<sup>‡</sup>Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico; <sup>§</sup>International Institute for Complex Adaptive Matter, University of California at Davis, Davis, California; <sup>¶</sup>Department of Chemistry and Department of Physics, State University of New York at Stony Brook, Stony Brook, New York; <sup>||</sup>State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin, China; and <sup>\*\*</sup>Center for Theoretical Biological Physics and Department of Physics, University of California at San Diego, San Diego, California

**ABSTRACT** We present a method for calculating the configurational-dependent diffusion coefficient of a globular protein as a function of the global folding process. Using a coarse-grained structure-based model, we determined the diffusion coefficient, in reaction coordinate space, as a function of the fraction of native contacts formed  $Q$  for the cold shock protein ( $TmCSP$ ). We find nonmonotonic behavior for the diffusion coefficient, with high values for the folded and unfolded ensembles and a lower range of values in the transition state ensemble. We also characterized the folding landscape associated with an energetically frustrated variant of the model. We find that a low-level of frustration can actually stabilize the native ensemble and increase the associated diffusion coefficient. These findings can be understood from a mechanistic standpoint, in that the transition state ensemble has a more homogeneous structural content when frustration is present. Additionally, these findings are consistent with earlier calculations based on lattice models of protein folding and more recent single-molecule fluorescence measurements.

## INTRODUCTION

The energy landscape theory of protein folding (1–5) has been an invaluable theoretical framework for understanding protein folding (6–10), oligomerization (11–13), and functional transitions (14–18). According to the theory, the energy landscape associated with protein folding lacks large energetic traps and has an overall funnel shape where the native ensemble is the lowest energy state. These minimally frustrated landscapes can be idealized as being devoid of energetic roughness, which enables the use of structure-based (Gō-like) models (8,10,19–23) to study the thermodynamic and kinetic properties of the folding process. Because these structure-based models lack energetic trapping, they also provide a means to characterize the topological contributions to folding.

Although there is a strong correlation between simulated barrier heights and experimental folding times (24), rates are a consequence of both the free-energy profile and the diffusion coefficient (25,26). Accordingly, direct comparison between experiments and theory requires both quantities. In principle, one may circumvent the need for the diffusion coefficient by simulating many thousands of folding trajectories and calculating the mean first passage time of folding (27–29). Such approaches are often computationally intractable and they do not always advance our physical understanding of the process. Therefore, it is desirable to

calculate both the free energy  $F$  and the diffusion coefficient  $D$  as functions of a global folding coordinate  $Q$  (30–34). If  $D$  is constant, then it only serves as a prefactor to the folding rate. However, when  $D$  is not constant, as we describe below, it can give rise to kinetic barriers in addition to the thermodynamic barriers (32).

The diffusion coefficient  $D$  is a result of the underlying energy landscape. As every conformation has a unique set of locally accessible interactions,  $D$  is a function  $Q$ . Although  $Q$  can be defined by a variety of measures, here we use the fraction of native contacts, as it has been shown to capture, accurately, the transition state ensemble of two-state proteins (35). Low values of  $Q$  correspond to the unfolded state and high values correspond to the folded ensemble. When  $Q$  is low, energetic contributions are largely from water-protein interactions. In the folded state (high  $Q$ ), the burial of hydrophobic surface area can be the dominant energetic contribution. In these two regimes, the local energetic roughness can be quite different, which can lead to different diffusion coefficients.

Many recent efforts have attempted to characterize  $D(Q)$  via experimental methods (36–42) and theoretical calculations (26,32–34,43–49). These studies have found that diffusion is not constant as a protein folds to the native state. This naturally leads to the question: Does diffusion vary with the degree of compactness because of energetic trapping, or topological constraints? To address this, we calculate  $D$  for a structure-based model that lacks energetic roughness and compare the findings to variants of the model that include energetic frustration.

In this article, we present the diffusive properties of a  $C_\alpha$  structure-based model in molecular dynamics simulations.

Submitted November 19, 2009, and accepted for publication April 14, 2010.

<sup>△</sup>Ronaldo J. Oliveira and Paul C. Whitford contributed equally to this article.

\*Correspondence: [vleite@sjrp.unesp.br](mailto:vleite@sjrp.unesp.br) or [whitford@lanl.gov](mailto:whitford@lanl.gov)

Editor: Feng Gai.

© 2010 by the Biophysical Society  
0006-3495/10/07/0600/9 \$2.00

doi: 10.1016/j.bpj.2010.04.041

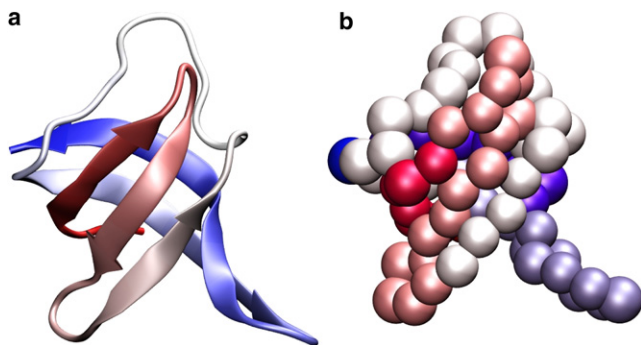


FIGURE 1 Cold-shock protein from *Thermotoga maritima* (*TmCSP* Protein DataBank entry 1G6P (36)), shown in (a) cartoon representation and (b)  $C_\alpha$  representation. The size of the atoms in panel b correspond to the excluded volume radii used in this model. The structures are colored from red (C-terminus) to blue (N-terminus) and were visualized with VMD (82). The *TmCSP* is a small globular protein with 66 amino acids, molecular mass of 7.5 kDa, and a three-dimensional structure known as a Greek-key  $\beta$ -barrel (five  $\beta$ -strands divided in two antiparallel  $\beta$ -sheets).

We compare the results obtained from the unfrustrated model (i.e., only native interactions are stabilizing) with an energetically frustrated variant of the model, which allows for a quantitative, and qualitative, comparison of topological and energetic contributions to the diffusion coefficient. As there is a large body of experimental data available, including denaturant-dependent diffusion coefficient measurements, we chose to study the cold shock protein from the hyperthermophilic bacterium *Thermotoga maritima* known as *TmCSP* (36) (Fig. 1 a). *TmCSP* is a 66-amino-acid  $\beta$ -barrel protein that is known to have well-defined two-state folding behavior (50–53). Through comparison with previous computational, theoretical, and experimental results, we provide evidence of the degree of roughness present in *TmCSP*.

## MODELS AND METHODS

### Structure-based $C_\alpha$ model

Here, we employ a well-studied coarse-grained structure-based model (8). In this model, each residue is represented as a single bead, located at the position of the  $C_\alpha$  atom (Fig. 1 b). For unfrustrated simulations, only native interactions are stabilizing and all residue pairs not in contact in the native structure are given a repulsive interaction to prevent chain crossing. In this model, the native structure is the global energetic minimum and the landscape lacks energetic traps. Native contacts were determined by the Contact of Structural Units software package (54). The functional form of the potential is

$$\begin{aligned}
 V = & \sum_{\text{bonds}} \epsilon_r (r - r_o)^2 + \sum_{\text{angles}} \epsilon_\theta (\theta - \theta_o)^2 \\
 & + \sum_{\text{dihedrals}} \epsilon_\phi \left\{ [1 - \cos(\phi - \phi_o)] + \frac{1}{2} [1 - \cos(3(\phi - \phi_o))] \right\} \\
 & + \sum_{\text{contacts}} \epsilon_C \left[ 5 \left( \frac{\sigma_{ij}}{r} \right)^{12} - 6 \left( \frac{\sigma_{ij}}{r} \right)^{10} \right] + \sum_{\text{noncontacts}} \epsilon_{NN} \left( \frac{\sigma_{NN}}{r} \right)^{12},
 \end{aligned} \quad (1)$$

where  $\epsilon_r = 100$ ,  $\epsilon_\theta = 20$ ,  $\epsilon_\phi = 1$ ,  $\epsilon_C = 1$ ,  $\epsilon_{NN} = 1$ , and  $\sigma_{NN} = 4.0 \text{ \AA}$ .  $r_o$ ,  $\theta_o$ ,  $\phi_o$ , and  $\sigma_{ij}$  are given the values found in the native structure.

To model nonspecific energetic frustration, we introduced an additional attractive interaction between all residue pairs that are not in contact in the native state and are separated by at least four residues in sequence. The functional form of the nonnative interactions is

$$V_f(r) = -\epsilon_{NC} \exp \left\{ -\frac{(r - r_g)^2}{\sigma_g^2} \right\}, \quad (2)$$

with  $r_g = 6.5 \text{ \AA}$  and  $\sigma_g = 1.0 \text{ \AA}$ . The degree of energetic frustration is determined by  $\epsilon_{NC}$ . In this study, we performed simulations with  $\epsilon_{NC} = 0.1$ – $0.7$ .

### Biasing potential

To calculate the diffusion coefficient about a specific value of  $Q$ , we introduced umbrella potentials (55,56) that restrained each simulation to a specified range of  $Q$  values. See Supporting Material for technical details.

### $\phi$ -values analysis

Experimentally, the structural content of the transition state ensemble in proteins is often studied by measuring changes in native stability and folding/unfolding rates upon point mutations. An approximate kinetic measure of the protein structure around a mutated residue is given by (57,58)

$$\phi \equiv \frac{-RT \ln k_{\text{mut}} / k_{\text{wt}}}{\Delta \Delta G^0}, \quad (3)$$

where  $k_{\text{mut}}$  and  $k_{\text{wt}}$  are the mutant and the wild-type folding rates, and  $\Delta \Delta G^0$  is the change in stability of the folded state upon mutation.

From a simulation, one may also calculate  $\phi$ -values by determining the change in the thermodynamic free energy barrier upon site mutation and comparing it to the change in native stability  $\Delta \Delta G^{F-U}$ . Computationally, this is less demanding than trying to determine differences in folding rates upon mutation. The  $\phi$ -values from structure-based simulations for each native contact pair (residues  $i$  and  $j$ ) can be further approximated as (8,59)

$$\phi_{ij} = \frac{\Delta \Delta G^{TS-U}}{\Delta \Delta G^{F-U}} \approx \frac{P_{ij}^{TS} - P_{ij}^U}{P_{ij}^F - P_{ij}^U}, \quad (4)$$

where  $P_{ij}^X$  is the probability of a contact between  $i$  and  $j$  being formed in state  $X$  (with  $X$  being  $F$ ,  $TS$ , or  $U$ ). For ease of discussion, here, we report  $\phi_F$ -values averaged over all native contacts with residue  $i$ .

## RESULTS

### Diffusion coefficient is robust to changes in restraining potential

The primary objective of this study was to determine how the diffusion coefficient  $D$ , in reaction coordinate space, changes during the folding process of *TmCSP* (Fig. 1 a). To calculate  $D$ , we employed a  $C_\alpha$  structure-based model (Fig. 1 b) with a restraining potential to ensure that each simulation sampled the phase space local to a particular value of  $Q$  (see Models and Methods for full description). The restraining potential was harmonic, centered at  $Q^*$ , and was given a strength of  $K_Q$ . When adding such a restraint, one must first ensure that the quantities of interest are not dependent on the strength of the restraint. To ensure that the diffusion coefficients are

a result of the underlying energy landscape, and not the biasing potential, several sets of simulations were performed, each with a different strength of the restraint.

To calculate the diffusion coefficient from a simulation, we employed a quasiharmonic diffusive approximation (25)

$$D = \frac{\Delta Q(T)^2}{2\tau(T)}, \quad (5)$$

where  $\Delta Q(T)^2$  is the mean-squared fluctuations in  $Q$ , and  $\tau(T)$  is the relaxation time associated with the decay of the autocorrelation function of  $Q$ , i.e.,  $C_Q(t)$ . Here,  $Q(t)$  is defined as the fraction of native  $C_\alpha$ - $C_\alpha$  contacts formed as a function of time (see **Models and Methods**). To use Eq. 5, the value of  $K_Q$  must be in a range for which a quasiharmonic approximation is warranted and  $D$  is not dependent on  $K_Q$ .

To determine values of  $K_Q$  for which the quasiharmonic approximation is valid, we compared the probability distributions in  $Q$  for a variety of  $K_Q$  values. Fig. 2 *a* shows the probability distributions for several values of  $K_Q$  (where the harmonic restraint is centered at  $Q^* = 0.5$ ), each at the folding temperature in the unrestrained case. For  $K_Q = 10$ , the probability distribution is clearly bimodal, with one peak corresponding to nativelylike structures ( $Q \approx 0.8$ ) and one peak corresponding to the unfolded ensemble ( $Q \approx 0.2$ ). For  $K_Q = 50$ , the probability distribution possesses a single peak near the minimum of the restraining potential  $Q = 0.5$ . For  $K_Q = 100$ , the width of the distribution is further reduced. This additional reduction of the width is undesirable. Because the diffusion coefficient describes the multidimensional process of the protein escaping from local energetic/topological minima, an overly-strong restraint may lead to artifacts by disallowing some possible routes of escape. In that scenario, our calculations of  $D$  could probe the restraining potential and not the underlying energy landscape.

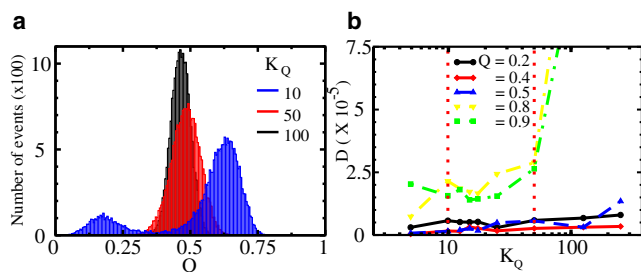


FIGURE 2 (a) Probability distributions in  $Q$  for biased ( $K_Q > 0$ ,  $Q^* = 0.5$ ) simulations with different strengths of the restraining potential  $K_Q$ . As  $K_Q$  increases, the distribution changes from a bimodal distribution, with peaks corresponding to the native and unfolded ensembles, to a single peak centered about the minimum of the restraining potential  $Q^*$ . The distribution is quasiharmonic for  $K_Q > 10$ . (b) The diffusion coefficient  $D$  is shown, on a semilog plot, as a function of  $K_Q$  for five values of  $Q^*$ . For  $10 < K_Q < 50$  (region delimited by the vertical dashed lines)  $D(Q)$  is relatively constant, demonstrating that estimates of  $D$  will be independent of  $K_Q$  over this interval. Simulations were performed at the folding temperature of the pure structure-based model  $T_f^0$ .

In addition to identifying a range of value of  $K_Q$  for which a quasiharmonic approximation is valid, we also determined a range of  $K_Q$ -values for which the calculated diffusion coefficients are not  $K_Q$ -dependent. Fig. 2 *b* shows  $D$  as a function of  $K_Q$  for a wide range of  $Q$  values. For low values of  $K_Q$  ( $< 10$ ), all calculated values of  $D$  increase with  $K_Q$ . As discussed above, this is due to the probability distribution being altered from a bimodal distribution to a distribution centered about the  $Q$  value of interest. For  $10 < K_Q < 50$  the calculated  $D$  is nearly constant for all  $Q$  values. Above  $K_Q = 50$ , the values of  $D$  again increase for  $Q > 0.7$ . Based on these data, we concluded that  $K_Q = 50$  will provide reliable values for the position-dependent diffusion.

### Diffusion coefficient dependence on $Q$

To understand the origins of the  $Q$ -dependence of the diffusion coefficient, one must consider the fluctuations in  $Q$  and the decay time of these fluctuations  $\tau_Q$ . As our calculated values of  $D$  are not sensitive to  $K_Q$  at  $K_Q \sim 50$ , all further values are reported for simulations performed at the folding temperature of the unrestrained simulations  $T_f^0$  with  $K_Q = 50$ . Fig. 3 *a* shows the time autocorrelation functions of  $Q$  for a variety of  $Q^*$  values. As  $Q^*$  is increased from 0.2 to 0.5 the characteristic decay time,  $\tau_Q$ , increases. At higher  $Q$  values ( $> 0.5$ ), the decay time decreases to a value smaller than in the unfolded ensemble (larger  $1/\tau_Q$  values in Fig. 3 *b*). The dispersion in  $Q$  ( $\Delta Q^2$ ) also displays a nonlinear dependence on  $Q$  (Fig. 3 *b*). Similar to  $\tau_Q$ ,  $\Delta Q^2$  initially increases with increasing  $Q$  (0.2–0.5) and then decreases as the native ensemble is reached ( $Q = 0.8$ ). Fluctuations in  $Q$ , shown in Fig. 3 *b*, rise considerably near the transition state due to the intrinsic instability of the transition state ensemble (TSE). In other words,  $Q$  exhibits large amplitude fluctuations as it overcomes the free energy barrier. For high and low  $Q$  values (the folded and unfolded ensembles),  $Q$  is

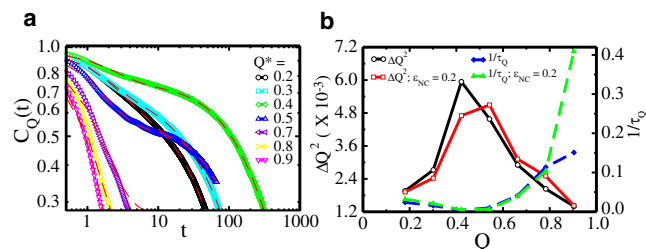


FIGURE 3 (a) Normalized correlation functions of  $Q$ ,  $C_Q(t)$ , shown on a log-log plot (time in reduced units) for different values of  $Q^*$ . Because a single exponential did not always fit well, each curve was fit to the sum of three exponentials to obtain an average decay time  $\tau_Q$ .  $K_Q = 50$  was used and the temperature was the folding temperature of the unbiased simulations  $T_f^0$ . The characteristic decay time  $\tau_Q$  is used to calculate the diffusion coefficient  $D$ . (b) The dispersion of the reaction coordinate  $\Delta Q^2$  as a function of the reaction coordinate  $Q$  (left axis), with and without energetic frustration. The inverse correlation time of  $Q$  ( $1/\tau_Q$ ) shown as a function of  $Q$  (right axis), with and without the frustration term. Calculations are shown at  $T_f$  of each  $\epsilon_{NC}$ .

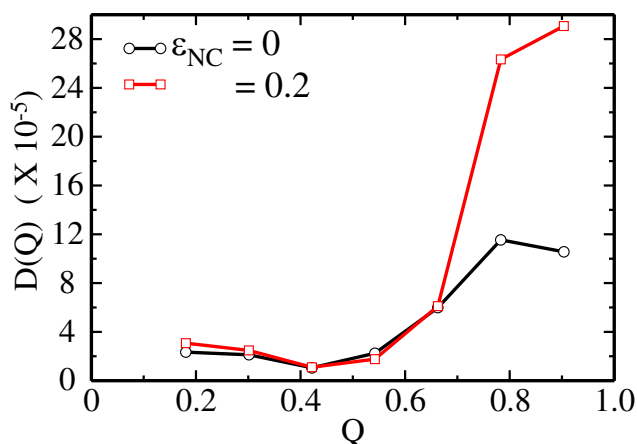


FIGURE 4 Comparison of the diffusion coefficients calculated without energetic frustration (black circles) and with energetic frustration ( $\epsilon_{NC} = 0.2$ , red squares) as functions of  $Q$ . All results were calculated with a restraining potential of strength  $K_Q = 50$  and at the  $T_f$  of each  $\epsilon_{NC}$ .

highly localized, which results in a small dispersion in  $Q$ . While  $\Delta Q^2$  is roughly symmetric about the TSE, the autocorrelation time, as well as its inverse, is asymmetric. This symmetry-breaking in  $\tau_Q$  leads to an asymmetric relationship between  $D$  and  $Q$ .

The diffusion coefficient as a function of  $Q$ ,  $D(Q)$ , is proportional to the product of  $\Delta Q^2$  and  $1/\tau_Q$ , and is shown in Fig. 4. We find  $D$  has large variations as a function of the folding reaction, which is in agreement with earlier studies on lattice models (32) and analytic studies (30,31,33,43). As discussed above,  $D(Q)$  (Fig. 4) largely follows  $1/\tau_Q$  (Fig. 3 b). Fluctuations of the reaction coordinate  $\Delta Q^2$  appear to have less influence on  $D(Q)$ , as  $\Delta Q^2$  changes only modestly with  $Q$ . The fact that  $D(Q)$  reaches a minimum around the TSE suggests the presence of a kinetic barrier, in addition to a thermodynamic one. After the protein moves from the TSE to the folded state,  $D(Q)$  once again increases and eventually reaches values that are 10-times larger than those corresponding to the unfolded ensemble.

The one-dimensional position-dependent diffusion coefficient variations indicate that the ruggedness of the energy landscape is not the same over the one-dimensional configuration space.  $D(Q)$  describes the local moves over microscopic barriers that connect states with similar values of  $Q$ . If the microstate is deep, it acts like a speed bump slowing both the drift and the superimposed Brownian movement (60) (i.e., the diffusion coefficient becomes small and escape-time from traps increases (7)). Because our energetically unfrustrated model gives rise to values of  $D$  that vary with  $Q$ , our results clearly indicate that the topology of the ensemble about a particular value of  $Q$  is inextricably linked to the diffusive dynamics. In other words, each configuration of the protein has a particular set of accessible escape routes, independent of the energetic roughness, which lead to the nonconstant form of  $D(Q)$ .

## Energetic frustration alters the diffusive dynamics

Due to the funnel-like nature of protein-folding energy landscapes, completely unfrustrated models, such as the one employed in this study, are sufficient to capture many aspects of protein folding (7,9,59,62–73). However, there is mounting evidence that a low degree of frustration can lead to accelerated folding rates (63) and provide a more accurate description of the unfolded ensemble (74). Such findings suggest a potential influence of energetic frustration on the diffusive properties associated with protein folding. To investigate this further, we employed a modified structure-based model in which the degree of frustration may be controlled. Specifically, we used the structure-based  $C_\alpha$ -model and added nonspecific attractive interactions between all nonnative atom pairs, where the functional form is a Gaussian with an energetic weight  $\epsilon_{NC}$  (see Models and Methods). Accordingly,  $\epsilon_{NC} = 0$  corresponds to the purely structure-based model.

Thermodynamic quantities were calculated for each frustrated system ( $\epsilon_{NC} > 0.0$ ) with  $K_Q = 0$ . For each parameter set, the fraction of native proteins  $f_N(T)$  was defined as

$$f_N(T) = \frac{\int_{\text{native}} \exp[-F(Q)/k_B T] dQ}{\int_0^1 \exp[-F(Q)/k_B T] dQ}, \quad (6)$$

where  $F(Q)$  is the free energy as a function of  $Q$ , the integral in the numerator is over all native conformations, and the denominator is over all possible  $Q$  values. We define the folding temperature  $T_f$  as the temperature where  $f_N = 0.5$  (dotted horizontal line in Fig. 5 a). As the degree of frustration is increased from 0,  $T_f$  initially increases and reaches its maximum at  $\epsilon_{NC} = 0.2$ . Because  $T_f$  measures thermodynamic stability, an increase in native-state stability with increased nonnative interaction strength may be surprising. This feature has two origins. First, in the native state ensemble, proteins are constantly fluctuating (75), which allows nonnative residue pairs to fluctuate toward and away from each other and form transient nonnative interactions (10). When nonspecific interactions are stabilizing, these transient nonnative interactions increase the stability of near-native conformations. The second contribution to

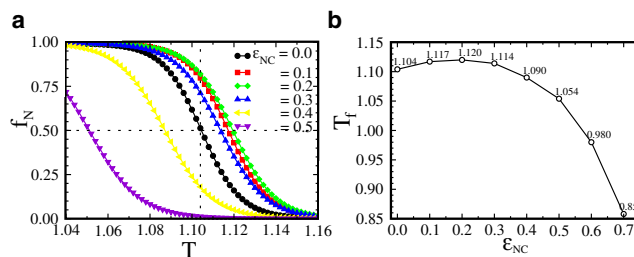


FIGURE 5 (a) Fraction of native protein  $f_N$  versus temperature  $T$  for different degrees of frustration  $\epsilon_{NC}$ . The value  $f_N$  was calculated using Eq. 6. (b) Folding temperature  $T_f$  as a function of  $\epsilon_{NC}$  for  $K_Q = 0$ . The folding temperature  $T_f$  has a maximum at  $\epsilon_{NC} = 0.2$ .

the increased native-state stability may be due to using a  $C_\alpha$  representation. All nonnative interactions were given energetic minima at 6.5 Å. When coarse-graining,  $C_\alpha$  pairs may be within that distance, but the side-chain configurations may lead to these pairs being considered not-in-contact. Thus, the noncontacting residues may stabilize these native configurations via the nonspecific interactions. Above  $\epsilon_{NC} = 0.2$ , the energetic frustration stabilizes the unfolded ensemble more than the folded ensemble and the folding temperature decreases, as expected. These findings are consistent with experimental results indicating that weakly attractive nonspecific interactions can increase the stability of Src homology 3 domain (70). In addition to affecting native stability, these experiments also revealed variations in the thermodynamic properties of the transition state ensemble, which were manifested as increased unfolding and refolding rates.

Increased levels of frustration also have direct effects on the calculated  $D(Q)$  profiles (Fig. 4). Similar to the unfrustrated simulations, the majority of the changes in  $D(Q)$  may be attributed to fluctuations in  $\tau_Q$  (Fig. 3). When energetic frustration is introduced,  $\Delta Q^2$  is only marginally perturbed while  $1/\tau_Q$  exhibits substantial deviations (Fig. 3). Comparison of the frustrated and unfrustrated simulations (Fig. 3) indicates that frustration has little effect on the  $\Delta Q^2$  and  $1/\tau_Q$  values associated with the unfolded ensemble ( $Q < 0.5$ ). After the protein has reached the folding transition state ( $Q \approx 0.5$ ) and moves to higher  $Q$  values,  $1/\tau_Q$  increases for both the unfrustrated and  $\epsilon_{NC} = 0.2$  simulations, although there is a larger increase in  $1/\tau_Q$  for the frustrated simulations than the unfrustrated ones. This finding may be counterintuitive, but it shows that a low degree of frustration can actually reduce the height of the microscopic barriers that are described by the diffusion coefficient.

### Residual entropy of the native state ensemble

The fact that  $D(Q)$  reaches a maximum in the folded ensemble can be understood by analyzing the density of states as a function of  $Q$ . Fig. 6 shows the density of states as a function  $Q$  for a lattice model (32) and the presented  $C_\alpha$  model. Although there is an increase in  $D(Q)$  as  $Q$  goes to 1,  $D(Q)$  remains on the same scale for large  $Q$  as for small  $Q$ . This is due to the ensemble nature of the native state. That is, in the  $C_\alpha$  model, the protein may interconvert between local structures without changing the value of  $Q$ , even when all native contacts are formed ( $Q = 1$ ). This leads to a degenerate native state, residual entropy, and nonzero correlation times. In contrast, in the lattice model, every possible move from the  $Q = 1$  state results in a decrease in  $Q$ . This nondegenerate native state leads to very low correlation times, and hence very large diffusion coefficients for the native state. Additionally, in the lattice model, single rearrangements can result in multiple contacts being formed or broken simultaneously. This lack of residual entropy in the

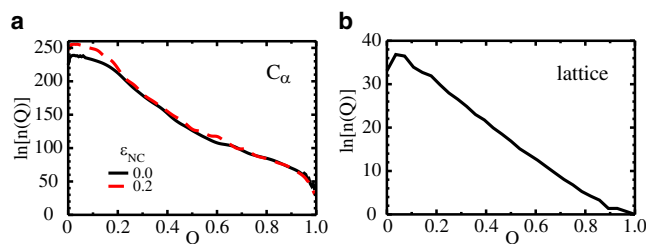


FIGURE 6 Density of states  $n$  as a function of  $Q$  for (a) the  $C_\alpha$  model and (b) a lattice model (32). High  $Q$  represents the folded ensemble and low  $Q$  is the unfolded ensemble. Panel a shows the density of states for the unfrustrated ( $\epsilon_{NC} = 0.0$ ) and frustrated ( $\epsilon_{NC} = 0.2$ ) systems. The  $C_\alpha$  model has a highly degenerate folded state, indicating the presence of residual entropy which allows for an increase in  $D$  after the protein passes through the TSE. For the lattice model there is a rapid decrease in the density of states with a nondegenerate folded state ( $n(E_{\text{folded}}) = n(Q_{\text{folded}}) = 1$ ) and monotonically decreasing values of  $D$  (32).

lattice model has prevented previous evaluation of  $D(Q = 1)$ . Here, by using an off-lattice  $C_\alpha$  model, we are able to calculate  $D(Q)$  for the full range of  $Q$ .

### Folding mechanism and $\phi$ -values analysis

The introduction of attractive nonnative interactions, or energetic frustration, changes the folding energy landscape (62,63,76) and can alter the structural content of the transition state ensemble. The  $\epsilon$ -values are commonly used experimentally to measure the degree of native structural content in the TSE about each residue. Computationally,  $\phi$ -values can also be determined, where a value of 0 indicates no native structural content and 1 indicates full structural content in the TSE. Fig. 7 shows  $\phi_i$  ( $\phi$ -value for each residue  $i$ ) obtained from simulations with no frustration ( $\epsilon_{NC} = 0.0$ ) and a low degree of frustration ( $\epsilon_{NC} = 0.2$ ). When  $\epsilon_{NC} = 0.2$ ,

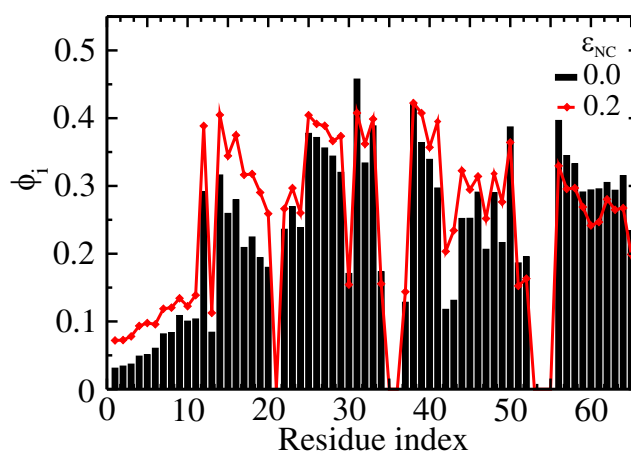


FIGURE 7 The  $\phi$ -values calculated for the unfrustrated simulations ( $\epsilon_{NC} = 0.0$ , black bars) and a weakly-frustrated system ( $\epsilon_{NC} = 0.2$ , red diamonds). Many residues with lower  $\phi_i$ -values increase, and high values decrease, upon the addition of frustration. This indicates a more homogeneous TSE when a low level of frustration is introduced. Simulations were performed without a restraining potential ( $K_Q = 0$ ) and at  $T_f(\epsilon_{NC})$ .

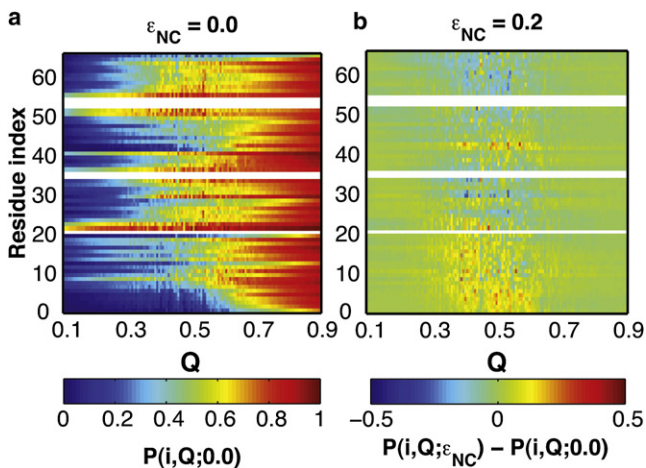


FIGURE 8 Probability of contacts being formed for each residue  $P(i, Q, \epsilon_{NC})$  as a function of the reaction coordinate  $Q$ , for the unfrustrated and weakly frustrated system. (a)  $P(i, Q, 0)$  increases (blue to red) as the protein folds. (b) Difference between the probabilities for the frustrated and unfrustrated simulations,  $P(i, Q, \epsilon_{NC}) - P(i, Q, 0.0)$ , where  $\epsilon_{NC} = 0.2$ . The simulations were performed at  $T = T_f$ . (Open bars) Residues lacking any native contacts.

there is a shift in structural content from the C-terminus to the N-terminus. Specifically, when  $\epsilon_{NC}$  is increased from 0.0 to 0.2 the  $\phi$ -values for residues 0–20 increase and residues 55–65 decrease. Additionally, other regions with low  $\phi$ -values, such as residues 42–45, increase when frustration is introduced.

Overall, introducing a low-degree of frustration appears to (partially) homogenize the  $\phi$ -values. Because residues that are less native (low  $\phi$ -values) are not surrounded by formed native interactions, they are more exposed to nonnative interactions. Accordingly, nonspecific stabilizing interactions are more accessible to less-native residues than highly-native ones. These nonspecific attractive interactions can then localize the residues involved in native-contacts, which results in additional native structure formation and a shift in the  $\phi$ -values.

To characterize the effects of frustration on the unfolded and folded basins, in addition to the TSE, we calculated the probability of contacts being formed with each residue, as functions of  $Q$  (Fig. 8). The probabilities for the unfrustrated case are shown in Fig. 8 a. For low  $Q$ , the probabilities are not homogeneously distributed, but are high around residues 22, 53, and 56, and nearly zero for all other residues. At approximately the transition state ( $Q = 0.5$ ), the probabilities follow the  $\phi$ -values, and have peaks around residues 22, 35, and 55. After passing the transition state, these regions may be considered nucleation sites, about which the rest of the protein's native structure is formed. Fig. 8 b shows the changes in the probabilities when frustration is introduced. Blue corresponds to decreased structure and red indicates increased structure formation. Similar to the  $\phi$ -values, in the TSE there is a shift in probabilities from the C-terminal

residues to the N-terminal residues. Surprisingly, the effects of the energetic frustration appear to be isolated to the TSE. One explanation for this feature is that frustration in the unfolded ensemble may not be well described by Gaussian potentials, as we have employed here. Instead, longer-range, screened-electrostatic interactions may be a larger contributor to frustration in the unfolded ensemble (74). In contrast, using a coarse-grained structure-based model, Das et al. (77) showed that introducing nonnative interactions and energetic heterogeneity has a large effect on the TSE, and improves agreement between experimental and theoretical  $\phi$ -values for Src homology 3 domain. Although our finding suggest short-range frustration is most important in the TSE, real proteins likely exhibit a combination of short-range and long-range nonnative interactions. Further investigation will be necessary to untangle the relationship among different types of frustration, the folding mechanism, and the diffusive dynamics of the folding process.

Perl et al. (78) explored the role of the chain termini residues on the folding stability by comparing the cold shock proteins *BcCSP* from the thermophile *Bacillus caldolyticus* with its homolog *BsCSPB* *Bacillus subtilis*. These two cold shock proteins have natively activated states of folding, similar to that of the hyperthermophilic *Thermotoga maritima* *TmCSP* (50) studied in this work. Their studies illustrate that major contributors to the difference in stability are residue 3 (which takes on natively structure in the TSE) and the C-terminal residue 66 (which forms late in the folding process) (78). Despite the fact that the C-terminal residues have high  $\phi$ -values in experiments and low  $\phi$ -values for the unfrustrated model, as discussed above, the  $\phi$ -values of the termini increase with increased energetic roughness. As suggested by our analysis of stability as a function of roughness, this comparison also demonstrates that cold shock protein likely has a modest degree of energetic roughness, though the exact degree, and type, of frustration cannot be unambiguously determined from the presented simulations.

## CONCLUSIONS

In this work, we have studied the folding of *TmCSP* using a coarse-grained structure-based model and we calculated the diffusion coefficient as a function of a reaction coordinate  $Q$ . Our main results can be outlined as follows: The diffusion coefficient displays nonmonotonic behavior as a function of  $Q$ , which can be attributed to a residual entropy of the native state ensemble. A role of residual entropy has been suggested previously (33), though here we explicitly calculate it and show its relationship to the diffusive dynamics. By introducing varied degrees of energetic roughness, we have shown that for low levels of frustration, *TmCSP* displays increased thermal stability and diffusion coefficients, relative to the unfrustrated regime, which agrees with previous findings (62,63). As frustration is increased, the stability reaches

a maximum, after which increased frustration leads to a less stable protein. At this optimum degree of frustration, the transition state is characterized by a more homogeneous distribution of  $\phi$ -values, relative to the unfrustrated case. In addition to changes in the structural content of the TSE, the diffusion coefficient is also affected by a low-degree of energetic frustration.

This work has shown that the diffusive dynamics are intimately linked to the topological and energetic aspects of a protein, and lays a foundation for understanding the diffusive properties of protein folding. Many examples can be found where the diffusion coefficient provides a nontrivial contribution to the folding dynamics. For example, as the presence of additional small free-energy barriers can actually accelerate folding rates (79,80), there must be a balance between folding barriers and diffusion along the reaction coordinate. The folding of proteins with smaller free-energy barriers, such as BBL (81), will also depend more on the precise structure of the diffusion coefficient (44,41,42), such that the nonmonotonic behavior of the diffusion may be the limiting factor that determines folding rates. With the presented framework, further investigation will explore the details of how the diffusive dynamics contributes to the folding of these and other systems.

## SUPPORTING MATERIAL

Additional text and one reference is available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(10\)00539-4](http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00539-4).

We thank the Center for Theoretical Biological Physics for use of its computing facilities.

This work was supported by the Center for Theoretical Biological Physics sponsored by the National Science Foundation (grant No. PHY-0822283, with additional support from NSF grant No. MCB-0543906). R.J.O., J.C., and V.B.P.L. were supported by Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. R.J.O. and J.C. have also been supported by the Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazilian Ministry of Education. P.C.W. thanks the U. S. National Science Foundation for its I2CAM International Materials Institute Award (grant No. DMR-0645461) to fund this international collaboration. J.W. thanks the National Science Foundation for its Career Award.

## REFERENCES

1. Leopold, P. E., M. Montal, and J. N. Onuchic. 1992. Protein folding funnels—a kinetic approach to the sequence structure relationship. *Proc. Natl. Acad. Sci. USA*. 18:8721–8725.
2. Bryngelson, J. D., J. N. Onuchic, ..., P. G. Wolynes. 1995. Funnel, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*. 21:167–195.
3. Bryngelson, J. D., and P. G. Wolynes. 1987. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA*. 84:7524–7528.
4. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
5. Wang, J., J. N. Onuchic, and P. G. Wolynes. 1996. Statistics of kinetic pathways on biased rough energy landscapes with applications to protein folding. *Phys. Rev. Lett.* 76:4861–4864.
6. Shoemaker, B. A., J. Wang, and P. G. Wolynes. 1997. Structural correlations in protein folding funnels. *Proc. Natl. Acad. Sci. USA*. 94:777–782.
7. Nymeyer, H., A. E. García, and J. N. Onuchic. 1998. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proc. Natl. Acad. Sci. USA*. 95:5921–5928.
8. Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural details of the transition state ensemble and “en-route” intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* 298:937–953.
9. Gosavi, S., L. L. Chavez, ..., J. N. Onuchic. 2006. Topological frustration and the folding of interleukin-1  $\beta$ . *J. Mol. Biol.* 357:986–996.
10. Whitford, P. C., J. K. Noel, ..., J. N. Onuchic. 2009. An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields. *Proteins Struct. Funct. Bioinf.* 75:430–441.
11. Levy, Y., and J. N. Onuchic. 2006. Mechanisms of protein assembly: lessons from minimalist models. *Acc. Chem. Res.* 39:135–142.
12. Yang, S. C., S. S. Cho, ..., J. N. Onuchic. 2004. Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. USA*. 101:13786–13791.
13. Lu, Q., H. P. Lu, and J. Wang. 2007. Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. *Phys. Rev. Lett.* 98:128105.
14. Whitford, P. C., O. Miyashita, ..., J. N. Onuchic. 2007. Conformational transitions of adenylate kinase: switching by cracking. *J. Mol. Biol.* 366:1661–1671.
15. Schug, A., P. C. Whitford, ..., J. N. Onuchic. 2007. Mutations as traps: two competing native conformations of the Rop-dimer. *Proc. Natl. Acad. Sci. USA*. 104:17674–17679.
16. Best, R. B., Y. G. Chen, and G. Hummer. 2005. Slow protein conformational dynamics from multiple experimental structures: the helix/sheet transition of arc repressor. *Structure*. 13:1755–1763.
17. Zuckerman, D. M. 2004. Simulation of an ensemble of conformational transitions in a united-residue model of calmodulin. *J. Phys. Chem. B*. 108:5127–5137.
18. Lu, Q., and J. Wang. 2008. Single molecule conformational dynamics of adenylate kinase: energy landscape, structural correlations, and transition state ensembles. *J. Am. Chem. Soc.* 130:4772–4783.
19. Pincus, D. L., S. S. Cho, ..., D. Thirumalai. 2008. Minimal models for proteins and RNA: from folding to function. In *Molecular Biology of Protein Folding*, Pt B, Vol. 84, Progress in Molecular Biology and Translational Science.. Elsevier/Academic Press, Amsterdam, The Netherlands.
20. Mickler, M., R. I. Dima, ..., M. Rief. 2007. Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. *Proc. Natl. Acad. Sci. USA*. 104:20268–20273.
21. Hills, Jr., R. D., and C. L. Brooks, III. 2009. Insights from coarse-grained Gō models for protein folding and dynamics. *Int. J. Mol. Sci.* 10:889–905.
22. Karanicolas, J., and C. L. Brooks, 3rd. 2003. Improved Gō-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. *J. Mol. Biol.* 334:309–325.
23. Ueda, Y., H. Taketomi, and N. Gō. 1975. Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effects of specific amino acid sequence represented by specific inter-unit interactions. *Int. J. Peptide Res.* 7:445–459.
24. Chavez, L. L., J. N. Onuchic, and C. Clementi. 2004. Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates. *J. Am. Chem. Soc.* 126:8426–8432.
25. Soccì, 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.

26. Baumketner, A., and Y. Hiwatari. 2002. Diffusive dynamics of protein folding studied by molecular dynamics simulations of an off-lattice model. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 66:011905.
27. Pande, V. S., A. Y. Grosberg, and T. Tanaka. 1997. On the theory of folding kinetics for short proteins. *Fold. Des.* 2:109–114.
28. Du, R., V. S. Pande, ..., E. S. Shakhnovich. 1998. On the transition coordinate for protein folding. *J. Chem. Phys.* 108:334–350.
29. Pande, V. S., and D. S. Rokhsar. 1999. Molecular dynamics simulations of unfolding and refolding of a  $\beta$ -hairpin fragment of protein G. *Proc. Natl. Acad. Sci. USA.* 96:9062–9067.
30. 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.
31. Lee, C. L., C. T. Lin, ..., 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 Stat. Nonlin. Soft Matter Phys.* 67:041905.
32. Chahine, J., R. J. Oliveira, ..., J. Wang. 2007. Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *Proc. Natl. Acad. Sci. USA.* 104:14646–14651.
33. Yang, S., J. N. Onuchic, and H. Levine. 2006. Effective stochastic dynamics on a protein folding energy landscape. *J. Chem. Phys.* 125:054910–054918.
34. Yang, S., J. N. Onuchic, ..., H. Levine. 2007. Folding time predictions from all-atom replica exchange simulations. *J. Mol. Biol.* 372:756–763.
35. Cho, S. S., Y. Levy, and P. G. Wolynes. 2006. P versus Q: structural reaction coordinates capture protein folding on smooth landscapes. *Proc. Natl. Acad. Sci. USA.* 103:586–591.
36. Kremer, W., B. Schuler, ..., H. R. Kalbitzer. 2001. Solution NMR structure of the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima*. *Eur. J. Biochem.* 268:2527–2539.
37. Nettels, D., I. V. Gopich, ..., B. Schuler. 2007. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA.* 104:2655–2660.
38. Hoffmann, A., A. Kane, ..., B. Schuler. 2007. Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA.* 104:105–110.
39. Nettels, D., I. V. Gopich, ..., B. Schuler. 2008. Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J. Phys. Chem.* 112:6137–6146.
40. Karplus, M., and D. L. Weaver. 1994. Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci.* 3:650–668.
41. Kubelka, J., T. K. Chiu, ..., J. Hofrichter. 2006. Sub-microsecond protein folding. *J. Mol. Biol.* 359:546–553.
42. Gruebele, M. 2005. Downhill protein folding: evolution meets physics. *C. R. Biol.* 328:701–712.
43. 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.
44. Kubelka, J., J. Hofrichter, and W. A. Eaton. 2004. The protein folding ‘speed limit’. *Curr. Opin. Struct. Biol.* 14:76–88.
45. Oliveberg, M., and P. G. Wolynes. 2005. The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* 38:245–288.
46. Hummer, G. 2005. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *N. J. Phys.* 7:34–48.
47. Best, R. B., and G. Hummer. 2010. Coordinate-dependent diffusion in protein folding. *Proc. Natl. Acad. Sci. USA.* 107:1088–1093.
48. Nettels, D., S. Müller-Späh, ..., B. Schuler. 2009. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc. Natl. Acad. Sci. USA.* 106:20740–20745.
49. Krivov, S. V., and M. Karplus. 2008. Diffusive reaction dynamics on invariant free energy profiles. *Proc. Natl. Acad. Sci. USA.* 105:13841–13846.
50. Perl, D., C. Welker, ..., F. X. Schmid. 1998. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat. Struct. Biol.* 5:229–235.
51. Wassenberg, D., C. Welker, and R. Jaenicke. 1999. Thermodynamics of the unfolding of the cold-shock protein from *Thermotoga maritima*. *J. Mol. Biol.* 289:187–193.
52. Schuler, B., W. Kremer, ..., R. Jaenicke. 2002. Role of entropy in protein thermostability: folding kinetics of a hyperthermophilic cold shock protein at high temperatures using  $^{19}\text{F}$  NMR. *Biochemistry.* 41:11670–11680.
53. Schuler, B., E. A. Lipman, and W. A. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 429:743–747.
54. Sobolev, V., R. Wade, ..., M. Edelman. 1996. Molecular docking using surface complementarity. *Proteins Struct. Funct. Genet.* 25:120–129.
55. Torrie, G. M., and J. P. Valleau. 1977. Nonphysical sampling distributions in Monte Carlo free-energy estimation—umbrella sampling. *J. Comput. Phys.* 23:187–199.
56. Bartels, C., and M. Karplus. 1997. Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations. *J. Comput. Phys.* 18:1450–1462.
57. Matouschek, A., J. T. Kellis, Jr., ..., A. R. Fersht. 1989. Mapping the transition state and pathway of protein folding by protein engineering. *Nature.* 340:122–126.
58. Fersht, A. R. 1995. Characterizing transition states in protein folding: an essential step in the puzzle. *Curr. Opin. Struct. Biol.* 5:79–84.
59. Nymeyer, H., N. D. Socci, and J. N. Onuchic. 2000. Landscape approaches for determining the ensemble of folding transition states: success and failure hinge on the degree of frustration. *Proc. Natl. Acad. Sci. USA.* 97:634–639.
60. Onuchic, J. N., Z. Luthey-Schulten, and P. G. Wolynes. 1997. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* 48:545–600.
61. Reference deleted in proof.
62. Plotkin, S. S. 2001. Speeding protein folding beyond the Gō model: how a little frustration sometimes helps. *Proteins Struct. Funct. Genet.* 45:337–345.
63. Clementi, C., and S. S. Plotkin. 2004. The effects of nonnative interactions on protein folding rates: theory and simulation. *Protein Sci.* 13:1750–1766.
64. Fan, K., J. Wang, and W. Wang. 2002. Folding of lattice protein chains with modified Gō potential. *Eur. Phys. J. B.* 30:381–391.
65. Li, L., L. A. Mirny, and E. I. Shakhnovich. 2000. Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus. *Nature.* 7:336–342.
66. Treptow, W. L., M. A. A. Barbosa, ..., A. F. P. de Araújo. 2002. Non-native interactions, effective contact order, and protein folding: a mutational investigation with the energetically frustrated hydrophobic model. *Proteins Struct. Funct. Bioinf.* 49:167–180.
67. Garcia, L. G., and A. F. P. de Araújo. 2006. Folding pathway dependence on energetic frustration and interaction heterogeneity for a three-dimensional hydrophobic protein model. *Proteins Struct. Funct. Bioinf.* 62:46–63.
68. Morton, V. L., C. T. Friel, ..., S. E. Radford. 2007. The effect of increasing the stability of non-native interactions on the folding landscape of the bacterial immunity protein Im9. *J. Mol. Biol.* 371:554–568.
69. Hamada, D., and Y. Goto. 1997. The equilibrium intermediate of  $\beta$ -lactoglobulin with non-native  $\alpha$ -helical structure. *J. Mol. Biol.* 269:479–487.
70. Viguera, A. R., C. Vega, and L. Serrano. 2002. Unspecific hydrophobic stabilization of folding transition states. *Proc. Natl. Acad. Sci. USA.* 99:5349–5354.



71. Di Nardo, A. A., D. M. Korzhnev, ..., A. R. Davidson. 2004. Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. *Proc. Natl. Acad. Sci. USA*. 101:7954–7959.
72. Neudecker, P., A. Zarrine-Afsar, ..., L. E. Kay. 2006. Identification of a collapsed intermediate with non-native long-range interactions on the folding pathway of a pair of Fyn SH3 domain mutants by NMR relaxation dispersion spectroscopy. *J. Mol. Biol.* 363:958–976.
73. Shan, B., D. Eliezer, and D. P. Raleigh. 2009. The unfolded state of the C-terminal domain of the ribosomal protein L9 contains both native and non-native structure. *Biochemistry*. 48:4707–4719.
74. Weinkam, P., E. V. Pletneva, ..., P. G. Wolynes. 2009. Electrostatic effects on funneled landscapes and structural diversity in denatured protein ensembles. *Proc. Natl. Acad. Sci. USA*. 106:1796–1801.
75. Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science*. 254:1598–1603.
76. Paci, E., M. Vendruscolo, and M. Karplus. 2002. Native and non-native interactions along protein folding and unfolding pathways. *Proteins Struct. Funct. Genet.* 47:379–392.
77. Das, P., S. Matysiak, and C. Clementi. 2005. Balancing energy and entropy: a minimalist model for the characterization of protein folding landscapes. *Proc. Natl. Acad. Sci. USA*. 102:10141–10146.
78. Perl, D., G. Holtermann, and F. X. Schmid. 2001. Role of the chain termini for the folding transition state of the cold shock protein. *Biochemistry*. 40:15501–15511.
79. Wagner, C., and T. Kiefhaber. 1999. Intermediates can accelerate protein folding. *Proc. Natl. Acad. Sci. USA*. 96:6716–6721.
80. Sinha, K. K., and J. B. Udgaonkar. 2008. Barrierless evolution of structure during the submillisecond refolding reaction of a small protein. *Proc. Natl. Acad. Sci. USA*. 105:7998–8003.
81. Cho, S. S., P. Weinkam, and P. G. Wolynes. 2008. Origins of barriers and barrierless folding in BBL. *Proc. Natl. Acad. Sci. USA*. 105: 118–123.
82. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14: 33–38, 27–28.