Excluded volume, local structural cooperativity, and the polymer physics of protein folding rates

Xianghong Qi and John J. Portman‡

Department of Physics, Kent State University, Kent, OH 44240

Edited by Peter G. Wolynes, University of California at San Diego, La Jolla, CA, and approved April 14, 2007 (received for review October 20, 2006)

A coarse-grained variational model is used to investigate the polymer dynamics of barrier crossing for a diverse set of two-state folding proteins. The model gives reliable folding rate predictions provided excluded volume terms that induce minor structural cooperativity are included in the interaction potential. In general, the cooperative folding routes have sharper interfaces between folded and unfolded regions of the folding nucleus and higher free energy barriers. The calculated free energy barriers are strongly correlated with native topology as characterized by contact order. Increasing the rigidity of the folding nucleus changes the local structure of the transition state ensemble nonuniformly across the set of proteins studied. Nevertheless, the calculated prefactors k_0 are found to be relatively uniform across the protein set, with variation in $1/k_0$ less than a factor of 5. This direct calculation justifies the common assumption that the prefactor is roughly the same for all small two-state folding proteins. Using the barrier heights obtained from the model and the best-fit monomer relaxation time 30 ns, we find that $1/k_0 \sim$ 1–5 μs (with average $1/k_0 \sim$ 4 μs). This model can be extended to study subtle aspects of folding such as the variation of the folding rate with stability or solvent viscosity and the onset of downhill folding.

nucleation | prefactor | topology

Folding in small proteins is often well characterized as a cooperative transition between two well defined structural populations: an unstructured globule ensemble and a structured folded ensemble. The transition rate between free energy minima is controlled by the dynamics of passing through an unstable transition region determined by saddlepoints in the free energy surface. Accordingly, the rate is expected to follow Arrhenius form

$$k_f = k_0 e^{-\beta \Delta F^\dagger}, \quad [1]$$

where $\beta = 1/k_B T$ is the inverse temperature and $\Delta F^\dagger$ is the free energy difference between the unfolded and transition-state ensembles. The exponential factor in Eq. **1** reflects the equilibrium population of the transition-state ensemble relative to unfolded ensemble and the prefactor, k_0, is the time scale associated with the dynamics of crossing the free energy barrier. Successful identification of specific residues structured in the transition-state ensemble by several different theoretical models (1–8) and numerous simulation studies (see, e.g., refs. 9–11 and references therein) has established that the topology of the native structure determines the folding mechanism of these proteins. In addition, two-state folding rates are well correlated with very simple measures of the native state topology such as contact order (12–15). While additive potentials often produce reasonable structural characterization of the transition state ensemble, the range of simulated folding rates and their relationship with contact order does not agree with experiment (10). In this paper we present direct folding rate calculations that capture the trends noted earlier in lattice models (16, 17) and very recently in continuum models (18, 19), leading several groups to speculate that the behavior of folding rates indicates enhanced structural cooperativity.

The term "structural cooperativity" usually refers to a mechanism by which the presence of a structured region makes additional order more favorable. For example, cooperativity is greater when a contact between two residues is more stabilized after one of the partners is already ordered. The additional stability is most naturally introduced through local attractive multibody interactions associated with coarse-grained potentials (18, 20, 21). Nonadditive potentials can also be neutral with respect to disordered and ordered residues and still increase cooperativity. For example, destabilizing partially ordered residues near structured resides generically sharpens the interface between folded and unfolded regions by increasing the surface energy of the folding nucleus (2, 22). Even purely repulsive interactions can enhance cooperativity. A good example of this is the "induced rigidity" (enhanced helical order) due to liquid crystal ordering in dense polymer solutions (23, 24) and globular helical proteins (25).

In the present analytic model, cooperativity is introduced through repulsive excluded-volume interactions between residues in proximity to native contact pairs. This potential is effectively "neutral" because it primarily destabilizes partially ordered residues at the interface of the folding nucleus. The cooperative term of the potential is pairwise additive in the space of all contacts, but it corresponds to an effective multibody potential when projected onto the set of native contacts. The particular form of cooperativity was developed so that the calculated barrier heights remain robust with respect to variations of excluded volume strength in the original variational model (3, 7).

Experimental evidence supporting a specific decomposition of the folding rate into the dynamic and thermodynamic factors in Eq. **1** is necessarily indirect (26). While structural predictions from models with a strong native state bias (Gō models) are robust, the value of the barrier height (and corresponding absolute time scale $1/k_0$) is more sensitive to details of the model (27, 28). The predicted prefactor is commonly assumed to be roughly uniform for different proteins with a magnitude of $O(0.1\text{–}1\ \mu s^{-1})$ (26, 29–31), although prefactors as large as $O(100\ \mu s^{-1})$ have been proposed recently (27, 32). While the precise value of the prefactor is a subdominant determinant of the absolute rate, accurate estimates give an important reference time scale essential, for example, to identify the fastest measured rates as downhill (or barrierless) folding (26, 33, 34). Calculations for 28 two-state proteins presented in this paper predict that the prefactor is relatively uniform on the order $O(1\ \mu s^{-1})$, largely independent of differences in the absolute folding rates

Author contributions: J.J.P. designed research; X.Q. and J.J.P. performed research; X.Q. and J.J.P. analyzed data; and X.Q. and J.J.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

‡To whom correspondence should be addressed. E-mail: jportman@kent.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0609321104/DC1.

or the native state topology. Furthermore, predicted folding rates agree with experimental trends provided interaction terms favoring modest structural cooperativity are included in the model. In particular, the relationship between barrier heights and contact order is found to be a consequence of relatively rigid folding nuclei.

Model: Excluded Volume and Cooperativity

The variational model developed by Portman, Takada, and Wolynes (7, 8) has proved reliable in predicting the structure of the transition-state ensemble of individual proteins at the residue level of resolution (3, 35, 36). In this model, the free energy of partially ordered ensembles of polymer configurations is developed through a reference Gaussian chain inhomogeneously constrained to the native positions by N harmonic variational constraints $\{C\}$. A summary of the variational model is given in the supporting information (SI) *Text*. Here, we focus on how enhanced cooperativity can be realized by the addition of repulsive interactions between nonnative contacts.

We divide the energy into two contributions

$$E[\{C\}] = \sum_{\text{NAT}} \langle u^{\text{nat}}(r_{ij})\rangle_0 + \sum_{\text{NON-NAT}} \langle u^{\text{coop}}(r_{ij})\rangle_0, \qquad [2]$$

where the subscript indicates an average over the reference Hamiltonian. The first term represents attractive interactions between monomers that are neighbors in the native structure (i.e., the Gō-model assumption). The second term, which is new to the model, represents excluded-volume interactions between nonnative contact pairs. Before explaining the consequences of these repulsive interactions, we first motivate the need for this contribution by considering the native contact potential in the original model.

The form of the interactions between native contacts is the sum of three Gaussians for convenience: $u^{\text{nat}}(r_{ij}) = \varepsilon_{ij} \sum_{a=l,i,s} \gamma_a e^{-\alpha_a r_{ij}^2}$, where ε_{ij} is the strength of the interaction (37). Repulsive intermediate-range and attractive long-range Gaussians sum to give a potential well with minimum $u^{\text{nat}}(r_{\min}) = -1$ at the distance $r_{\min} = 6$ Å. The short-range Gaussian represents excluded volume for native contact pairs; we choose γ_l for each contact to give the same strength at zero distance, $U(0)$. The finite strength of the repulsion at $r = 0$ is an artifact of the potential (and the finite native monomer density at short range), so there is some ambiguity in determining the appropriate value for $U(0)$. This is troubling because it was found that the calculated barrier height is sensitive to the value of $U(0)$, even though the structure of the transition state ensemble is relatively robust for most proteins.

The sensitivity of $\Delta F^{\dagger}/k_{\text{B}}T_{\text{f}}$ on the excluded volume strength $U(0)$ indicates that the cooperativity in the model is relatively low. This can be understood by considering the short-distance pair density of a partially ordered chain, $n_{ij}(\mathbf{r}) = \langle \delta(\mathbf{r} - \mathbf{r}_{ij})\rangle_0$. Integrating over angles gives the radial pair density

$$n_{ij}(r) \sim \frac{1}{\sqrt{a^2 \delta G_{ij}}} \frac{r}{s_{ij}} \sinh\left(\frac{3 r s_{ij}}{a^2 \delta G_{ij}}\right) \exp\left(-\frac{3}{2}\frac{r^2 + s_{ij}^2}{a^2 \delta G_{ij}}\right), \qquad [3]$$

where the correlations $G_{ij} = \langle \delta\mathbf{r}_i \cdot \delta\mathbf{r}_j \rangle_0 / a^2$ and $\delta G_{ij} = G_{ii} + G_{jj} - 2G_{ij}$ is the magnitude of the fluctuations about the relative mean separation $\mathbf{s}_{ij} = \sum (G_{ik} - G_{jk}) C_k \mathbf{r}_k^{\text{N}}$; and $a = 3.8$ Å is the distance between adjacent monomers. The weight at short distances ($r < r_0$, where r_0 is excluded volume interaction length scale) is small when the pair is sufficiently delocalized ($r_0 \ll a\sqrt{\delta G_{ij}}$) or sufficiently localized ($a\sqrt{\delta G_{ij}} \ll s_{ij}$). Consequently, the sensitivity of the barrier height on the short-distance repulsion is due to the partially ordered residues in the transition state ensemble. Increasing the cooperativity by destabilizing partially ordered residues makes the barrier height less sensitive to $U(0)$.

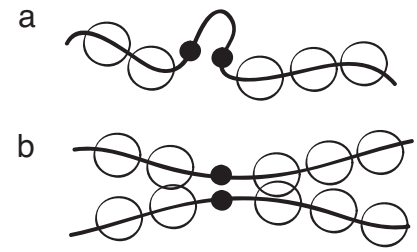

Fig. 1. Cooperativity of local and nonlocal contacts. (*a*) Localization of a short-ranged pair induces little cooperativity because the monomer density of the surrounding residues is not significantly altered by a partially formed local contact. (*b*) In contrast, if the contact pair is nonlocal sequences, this brings into proximity larger regions of residues (of the order of the persistence length), increasing the cooperativity. The dependence of the cooperativity on sequence length is reminiscent of cooperative desolvation between folded segments (60, 61).

To this end, we modify the interactions between native contacts through a repulsive potential between residues in close proximity to native contacts $u^{\text{coop}}(r_{ij})/\varepsilon_0 = U(0) \exp[-\alpha_s r_{ij}^2]$. This term increases cooperativity of partially structured ensembles by encouraging residues surrounding a native ordered pair to be either delocalized or ordered themselves to reduce short-distance density overlap.

When viewed as an effective potential involving only native contact pairs, the repulsion between nonnative contacts effectively induces local multibodied interactions. Because of chain connectivity, structured regions that are sufficiently nonlocal in sequence have greater cooperativity (see Fig. 1). Accordingly, we define a reduced set of nonnative contacts: for every native pair (i, j) with $|i - j| \geq 12$, we include pairs within a window $[i \pm 4, j \pm 4]$ and eliminate duplicates or native contact pairs from the sum. With this convention, the barrier height varies less than about 1–2 $k_{\text{B}}T_{\text{f}}$, over a wide range of $U(0)$ ($5 \leq U(0) \leq 60$). In the following we take $U(0) = 50$.

The parameters of the model are the same as given in ref. 7 except (*i*) the magnitude $U(0)$ is fixed for each contact; (*ii*) inclusion of the cooperativity term; (*iii*) the radius of gyration of the globule is set by the chain length according to the scaling law given in ref. 38. We note that with these parameters, the folding route for λ repressor studied in ref. 7 is structurally similar to the cooperative folding route, although the barrier is ~2 times larger.

Results

Folding Rates and Prefactors. We calculated the prefactors and folding routes of 28 two-state folding proteins. The corresponding folding rates at the transition midpoint, $k_{\text{f}} = k_0 e^{-\Delta F^{\dagger}/k_{\text{B}}T_{\text{f}}}$, are shown in Fig. 2. Calculated rates in absolute units depend on the time scale set by the monomer relaxation rate $\sigma = 3D_0/\text{a}^2$, which we take as a fitting parameter. As shown in Fig. 2, the predicted and measured rates are well correlated ($r = 0.8$) with agreement within an order of magnitude for 80% of the proteins.

The best-fit monomer relaxation time $1/\sigma = 30$ ns is on the order of the time scale of unfolding a helical segment (39). With this microscopic time scale, the longest relaxation time of a chain of 100 monomers is approximately $\tau_{\text{R}} \sim O(10\ \mu\text{s})$, which compares well (40) with the time scale for the fastest collapse kinetics measured in proteins and polypeptides (41–43). On the other hand, $1/\sigma$ is an order of magnitude slower than estimates obtained from an effective diffusion coefficient inferred from loop closure experiments of small peptides (~1 ns) and two orders of magnitude slower than estimates from bare diffusion coefficients of the monomer (~100 ps) (44). The source of small effective diffusion coefficients associated with simple Gaussian

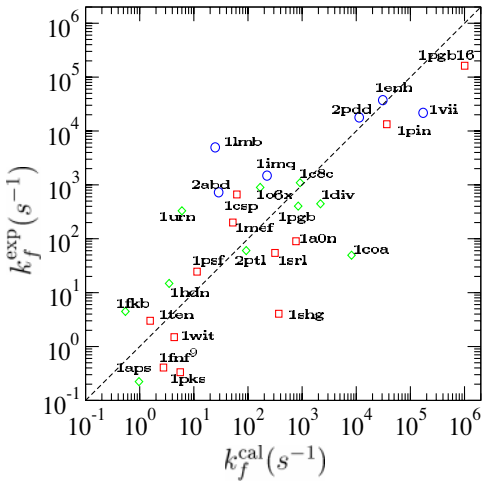

Fig. 2. Comparison between experimental (exp) and calculated (cal) folding rates. The correlation coefficient is $r = 0.8$, with $P = 4.4 \times 10^{-7}$. ○, α-protein; □, β-protein; ◇, $\alpha\beta$-protein. Kinetic data for proteins 1enh, 1vii, and 1hdn were taken from ref. 15; kinetic data for the rest of the proteins were taken from ref. 13. Here, proteins with structures determined by x-ray crystallography are 1pgb, 1coa, 1csp, 1c8c, 1pin, 1lmb, 1enh, 1fkb, 1urn, 1shg, 1ten, *1div, *1fnf9, and *1pgb16; proteins with structures determined by NMR are 1o6x, 1imq, 1srl, 2ptl, 1aps, 2abd, *1a0n, *2pdd, *1mef, *1psf, *1wit, *1pks, *1vii, and *1hdn. Asterisks indicate proteins not used in ϕ-value analysis.

models is not fully understood (45–47). Nevertheless, results from recent experiments on small peptides under different solvent conditions indicate that intrachain interactions (which can be interpreted kinetically as a kind of internal friction) induce local activation barriers that renormalize the effective monomer diffusion coefficient (48, 49). Although controversial, internal friction may explain why the speed limit of protein folding is fixed at $\sim O(0.1\text{–}1.0\ \mu s)$ (50).

Even though the prefactor of each protein is calculated individually, its value in absolute units ultimately depends on the calculated barrier heights through the microscopic time scale σ. Relative prefactors, on the other hand, are independent of this fitting parameter. As shown in Fig. 3, the distribution of prefactors is relatively uniform, varying within a factor of 5 for most proteins. Using the fitted value for σ, $\tau_0 = 1/k_0$ varies mainly between 1 and 5 μs, with an average $\bar{\tau}_0 = 4\ \mu$s. Given this narrow distribution, it is not surprising that a uniform prefactor of $\bar{k}_0 = (4\ \mu s)^{-1}$ gives essentially the same correlation to the measured

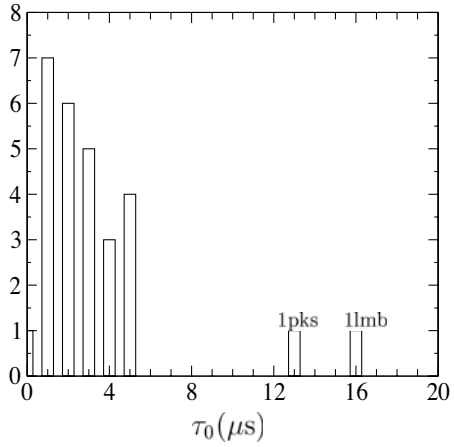

Fig. 3. Histogram of the inverse folding prefactor $\tau_0 = 1/k_0$ for the 28 two-state proteins. Ordinate is number of proteins.

and predicted rates (data not shown). Thus, direct calculation of the barrier-crossing dynamics gives solid evidence supporting the common assumption that the folding rate prefactor is largely independent of topology. Recent work by Henry and Eaton (28) also suggests the prefactor is relatively uniform across two-state folding proteins based on analysis of folding rates from a different set of analytic models. The value for the average prefactor $\bar{k}_0 \sim 10^5\ s^{-1}$ agrees within an order of magnitude with estimates based on semiempirical and theoretical models (5, 51, 52) as well as analysis of thermodynamic data from differential scanning calorimetry (53). This value also is consistent with the fastest measured rates, $\sim 1\ \mu$s, if the time scale for downhill folding is approximated by the Arrhenius rate with a vanishing barrier (26, 34).

A closer look at the two proteins (1lmb and 1pks) with exceptionally small calculated prefactors reveals that in each case the unstable mode becomes degenerate at a stability near the transition midpoint. The structure of the transition ensemble changes sharply, although continuously, as the temperature crosses the degenerate point. In particular, the curvature of the unstable mode (and consequently the calculated prefactor) sharply vanishes in a cusp catastrophe (54). Away from these isolated temperatures, the prefactors return to the range exhibited by the majority of the proteins studied. Several of the proteins studied have similar rapid changes of the transition state as a function of temperature, occurring at temperature sufficiently far away from the midpoint so that the prefactor is relatively unaffected near T_f. In this high-dimensional model, catastrophes can be generally expected when local minima and saddle-points merge at isolated values of the control parameters (e.g., temperature). The shape of the calculated prefactor versus temperature is thus determined by these degenerate points. For example, even for a route with a single transition state, the metastable unfolded or folded minimum disappears in a fold catastrophe at the limit of stability (spinodal) for both low and high temperatures (55, 56). If there are no other catastrophes, the calculated prefactor obtains a maximum at an intermediate temperature and vanishes at the spinodals. Near the maximum the prefactor varies much more slowly with temperature than near the spinodal. This generic shape of the prefactor is interesting because it can account qualitatively for nonlinear dependence of the rate with stability (chevron turnovers) and may indicate kinetic signatures anticipating the onset of downhill folding.

Nevertheless, interpreting these results requires some care. The harmonic expansion of the free energy is not expected to accurately reflect the global curvature of the free energy over $\sim F^{\dagger} \pm k_B T$ when the local curvature is very small. For these cases, it is likely that the formalism should be modified away from strictly local curvatures to get accurate estimates of the prefactor. Even if the renormalized prefactor is found to be relatively constant, the rapid change of the order parameter at the transition state that accompanies a catastrophe may itself account for chevron rollover, similar to the transition state switching mechanism suggested by Oliveberg and colleagues (57). The subtle variation of the prefactor and free energy barrier height with stability is an important issue that has yet to be thoroughly explored.

Free Energy Profiles and Folding Routes. In the context of identifying kinetic trends for all two-state proteins, a complete theory of the folding mechanism must reliably predict structural properties of the transition state ensemble in addition to absolute folding rates. The formation of local order along the folding route can be characterized by the degree of localization about the native positions

$$\rho_i[\{C\}] = \langle \exp[-3\alpha^N(\mathbf{r}_i - \mathbf{r}_i^N)^2/2a^2] \rangle_0. \quad [4]$$

BIOPHYSICS

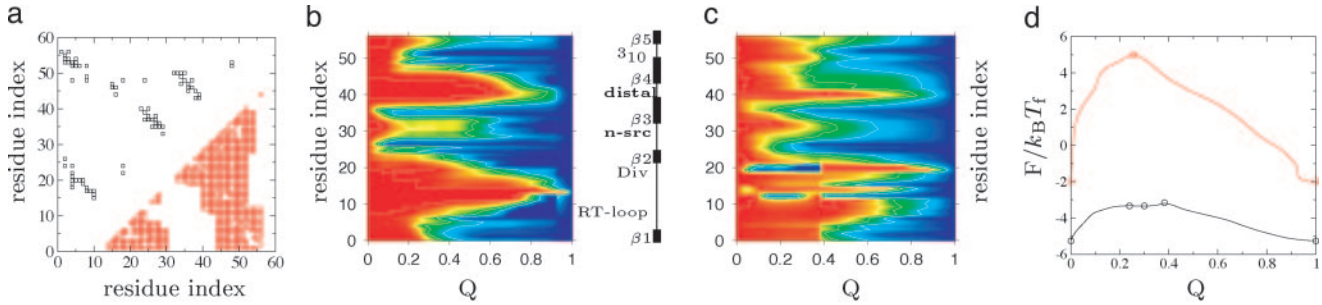

Fig. 4. Folding route for Src tyrosine kinase SH3 domain (Protein Data Bank ID code 1srl) characterized locally by the normalized native density $\tilde{\rho}_i = (\rho_i - \rho_i(G))/(\rho_i(N) - \rho_i(G))$ and the global progress coordinate $Q = 1/N \sum \tilde{\rho}_i$. In the center is a diagram of the protein secondary structure: a five-strand β-barrel (β1–β5) with three loops (RT, n-Src, distal), diverging turn and a 3_{10} helix. (*a*) Contact map: black squares are native contacts and red squares are nonnative pairs. (*b* and *c*) The local and global structure along the folding route with cooperativity (*b*) and without cooperativity (*c*). The degree of structural localization of each residue is reflected in the colors, linearly scaled from red ($\tilde{\rho}_i = 0$) to blue ($\tilde{\rho}_i = 1$). (*d*) Free energy profile as a function of Q: the red curve is the folding route with cooperativity, and the black curve is the route without cooperativity. The circles denote the critical points defining the folding route, and the curves are the steepest-descents paths. Experimental ϕ values suggest that β2–β3–β4 are ordered in the transition state ensemble (62), in agreement with the structure derived from the model.

We refer to ρ_i as the native density. Comparing the folding profiles and structural localization of the residues shown in Fig. 4 illustrates the cooperative nature of the folding routes induced by the repulsive nonnative interactions. While the coarse-grained structures of the transition state ensembles are similar for this protein, the residues order more gradually in the noncooperative routes. Still, even for the cooperative route, the interface has a finite width as the structural ensembles retain some partial ordering of the residues. The sharper interface of the cooperative route is also accompanied by a significantly larger barrier.

The effect of cooperativity on the structure of the transition state ensemble is complicated to describe in general. Cooperativity narrows the interface by destabilizing partially ordered residues in favor of either more ordered or more disordered. Whether a particular interfacial residue is excluded or incorporated into the folding nucleus is a subtle question, determined by the delicate balance between changes in entropy and energy due to localization. One measure to characterize changes in local structural order is the cross-correlation

$$\Omega = \hat{\rho}_{\text{coop}} \cdot \hat{\rho}_{\text{noncoop}}, \qquad [5]$$

where $\hat{\rho}_{\text{coop}}$ and $\hat{\rho}_{\text{noncoop}}$ denote unit vectors with elements $\rho_i[\{C\}]$ for transition state ensembles with and without cooperativity, respectively. Fig. 5 shows the value Ω for each protein as

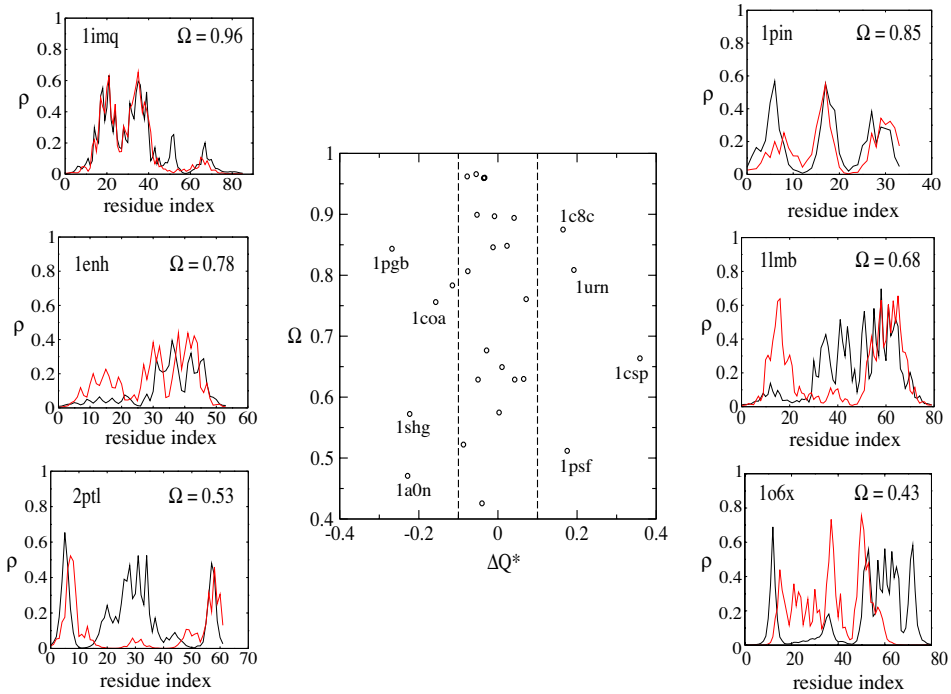

Fig. 5. Comparison of the structure of cooperative and noncooperative transition state ensembles. (*Center*) The change in the local structure of the transition state ensemble characterized by the cross-correlation coefficient Ω plotted against the change in the global order parameter ΔQ^*. (*Left* and *Right*) Typical examples of the native density profile for various values of overlap Ω. In each example, the red line corresponds to the cooperative route, and the black line corresponds to the noncooperative route. The corresponding correlation coefficients between measured and predicted ϕ are 1pin, 0.77; 1lmb, 0.84; 1enh, 0.87; 2ptl, 0.0; 1imq, −0.1; and 1o6x, −0.2.

well as a typical example of the overlap of native densities evaluated at Q^*. For 80% of the proteins studied, the overlap between the transition state ensemble structures is >60%. Nevertheless, the variation of Ω indicates this form of cooperativity does not effect every protein uniformly.

Changes of the transition state ensemble can be also be characterized by the variation of the global order parameter ΔQ^*. As shown in Fig. 5, the majority of the proteins studied have $|\Delta Q^*| \leq 0.1$. In terms of global order, the α-helical proteins are not very sensitive to cooperativity, although the local structure of the transition state ensemble can change significantly. For β and α/β proteins, some systematic errors in the calculated barrier height can be associated with relatively large changes in the global order. For proteins with $\Delta Q^* > 0.1$ (1urn, 1c8c, 1psf, 1csp), the model overestimates the barrier heights, whereas for proteins with $\Delta Q^* < 0.1$ (1pgb, 1a0n, 1coa, 1shg) the model underestimates the barrier height (see Fig. 2). This trend may be particular to the form of cooperativity used in this model.

Direct comparison between theoretical and measured ϕ values shows that cooperative routes generally have significantly higher correlation with experiment. Following Garbuzynskiy *et al.* (27), we make a distinction between contact maps obtained from native structures determined by x-ray crystallography and those from the first model of an NMR structure or minimized averaged NMR structure. Overall, the theory predicts ϕ values for studied x-ray structures reasonably well. Still, there are exceptions. Of the 11 x-ray structures (see legend of Fig. 2), two proteins(1shg and 1ten) have large negative correlations. The average correlation coefficient for 9 remaining proteins increases from 0.33 (for noncooperative routes) to 0.6 (cooperative routes). Predictions of ϕ values for NMR-determined structures is significantly worse, with the average 0.1 for both noncooperative routes and cooperative routes. In Fig. 5, we give three examples each for x-ray structures and NMR structures.

Folding Barriers and Absolute Contact Order. Because the prefactors are relatively uniform, the wide variation of relative folding rates is determined by differences in free energy barriers. To investigate the relationship between barrier heights and native topology, we consider the correlation between the free energy barriers and the absolute contact order (13). Fig. 6 shows that the calculated barrier height is highly correlated ($r = 0.9$) with absolute contact order when the cooperativity term is included in the model. The barrier heights calculated without cooperativity do not show significant correlation with absolute contact order. This observation indicates that the relationship between native topology and the folding rate is sensitive to the rigidity of the folding nucleus. This may in fact be a robust result, largely independent of the details of reasonable potentials that increase local cooperativity between native contacts (18).

Assuming the prefactor is roughly uniform, the range of measured rates for this protein set corresponds to a range of free energy barrier heights of $\sim 14 k_B T_f$, in agreement with the calculated barriers. In contrast, the range of barriers for the noncooperative routes spans only $\sim 5 k_B T_f$. Interestingly, this is the same range determined through coarse-grained Gō-model simulations (10, 58, 59). Furthermore, the low correlation between contact order and barrier heights of noncooperative routes is also reminiscent of results from Gō-model simulations (10). Together, these results suggest that the cooperativity of typical Gō-model simulations based on two-body pair potential is too low (16–19).

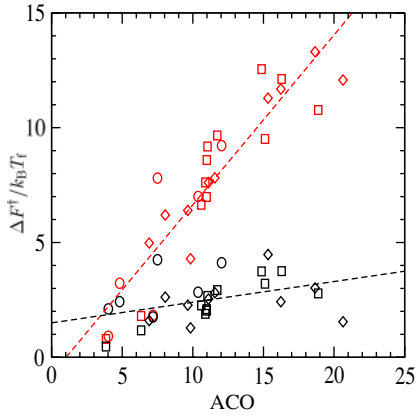

Fig. 6. Free energy barrier $F^{\ddagger}/k_B T_f$ plotted against the absolute contact order, $ACO = 1/N_{con} \sum |i - j|$ where the sum is over the N_{con} native contacts pairs. Red points correspond to cooperative folding routes, and black points correspond to noncooperative routes. Symbols have the same meaning as in Fig. 2. For the cooperative routes, the correlation coefficient is $r = 0.91$ with $P = 1.6 \cdot 10^{-11}$; for noncooperative routes, the correlation coefficient is $r = 0.41$ with $P = 0.03$.

Conclusion

The repulsive potential between residues in proximity to native contacts is a convenient way to alleviate sensitivity on the excluded volume strength in the original model. This approach was successful because the potential enhances cooperativity of the model. Our point of view is that the nature of the interface of the folding nucleus is key in determining the behavior of folding rates and mechanisms, regardless of the specific form of cooperative interactions or the microscopic origins. If the qualitative results from this study can be extended beyond this variation model, it is likely to be limited to models that enhance cooperativity locally. Because these results are robust with respect to the excluded volume strength $U(0)$, the model lacks flexibility to explore a wide range of surface tensions. It will be interesting to see whether these conclusions hold when the interfacial surface tension is controlled directly through, for example, the formalism of density functional theory of first-order nucleation.

This work was supported in part by grant awarded by the Ohio Board of Regents Research Challenge program.

1. Shoemaker BA, Wang J, Wolynes PG (1997) *Proc Natl Acad Sci USA* 94:777–782.
2. Shoemaker BA, Wang J, Wolynes PG (1999) *J Mol Biol* 287:675–694.
3. Portman JJ, Takada S, Wolynes PG (1998) *Phys Rev Lett* 81:5237–5240.
4. Alm E, Baker D (1999) *Proc Natl Acad Sci USA* 96:11305–11310.
5. Muñoz V, Eaton WA (1999) *Proc Natl Acad Sci USA* 96:11311–11316.
6. Galzitskaya OV, Finkelstein AV (1999) *Proc Natl Acad Sci USA* 96:11299–11304.
7. Portman JJ, Takada S, Wolynes PG (2001) *J Chem Phys* 114:5069–5081.
8. Portman JJ, Takada S, Wolynes PG (2001) *J Chem Phys* 114:5082–5096.
9. Clementi C, Nymeyer H, Onuchic JN (2000) *J Mol Biol* 298:937–953.
10. Koga N, Takada S (2001) *J Mol Biol* 313:171–180.
11. Ding F, Guo W, Dokholyan MV, Shakhnovich EI, Shea J-E (2005) *J Mol Biol* 350:1035–1050.
12. Plaxco KW, Simons KT, Baker D (1998) *J Mol Biol* 277:985–994.
13. Ivankov D, Garbuzynskiy SO, Alm E, Plaxco KW, Baker D, Finkelstein AV (2003) *Protein Sci* 12:2057–2062.
14. Makarov DE, Plaxco KW (2003) *Protein Sci* 12:17–26.
15. Bai Y, Zhou H, Zhou Y (2004) *Protein Sci* 13:1173–1181.
16. Jewett AI, Pande VS, Plaxco KW (2003) *J Mol Biol* 326:247–253.
17. Kaya H, Chan HS (2003) *J Mol Biol* 326:911–931.
18. Ejtehadi MR, Avall SP, Plotkin SS (2004) *Proc Natl Acad Sci USA* 101:15088–15093.
19. Kaya H, Liu Z, Chan HS (2005) *Biophys J* 89:520–535.
20. Eastwood MP, Wolynes PG (2001) *J Chem Phys* 114:4702–4716.
21. Wang J, Lee C, Stell G (2005) *Chem Phys* 316:53–60.
22. Wolynes PG (1997) *Proc Natl Acad Sci USA* 94:6170–6175.
23. Kim YH, Pincus P (1979) *Biopolymers* 18:2315–2322.

BIOPHYSICS

24. Flory PJ, Matheson RR (1984) *J Phys Chem* 88:6606–6612.
25. Luthey-Schulten ZA, Ramirez BE, Wolynes P (1995) *J Phys Chem* 99:2177–2185.
26. Yang WY, Gruebele M (2003) *Nature* 423:193–197.
27. Garbuzynskiy SO, Finkelstein AV, Galzitskaya OV (2004) *J Mol Biol* 336:509–525.
28. Henry ER, Eaton WA (2004) *Chem Phys* 307:163–185.
29. Li MS, Klimov DK, Thirumalai D (2004) *Polymer* 45:573–579.
30. Naganathan AN, Muñoz V (2005) *J Am Chem Soc* 127:480–481.
31. Kubelka J, Hofrichter J, Eaton WA (2004) *Curr Opin Struct Biol* 14:76–88.
32. Plotkin SS (2005) *Biophys J* 88:3762–3769.
33. Gruebele M (2005) *C R Biol* 328:701–712.
34. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J (2005) *J Mol Biol* 359:546–553.
35. Shen T, Hofmann CP, Oliveberg M, Wolynes PG (2005) *Biochemistry* 44:6433–6439.
36. Zong CH, Wilson CJ, Shen TY, Wolynes PG, Wittung-Stafshede P (2006) *Biochemistry* 45:6458–6466.
37. Miyazawa S, Jernigan RL (1996) *J Mol Biol* 256:623–644.
38. Jha AK, Colubri A, Freed KF, Sosnick TR (2005) *Proc Natl Acad Sci USA* 102:13099–13104.
39. Ivankov DN, Finkelstein AV (2001) *Biochemistry* 40:9957–9961.
40. de Gennes PG (1985) *J Phys (France)* 46:L639–L642.
41. Shastry MCR, Roder H (1998) *Nat Struct Biol* 5:385–392.
42. Qui L, Zachariah C, Hagen SJ (2003) *Phys Rev Lett* 90:168103:1–4.
43. Magg C, Schmid FX (2004) *J Mol Biol* 335:1309–1323.
44. Lapidus LJ, Eaton WA, Hofrichter J (1999) *Proc Natl Acad Sci USA* 97:7220–7225.
45. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) *J Phys Chem B* 106:11628–11640.
46. Yeh I-C, Hummer G (2002) *J Am Chem Soc* 124:6563–6568.
47. Portman JJ (2003) *J Chem Phys* 118:2381–2391.
48. Buscaglia M, Lapidus LJ, Eaton WA, Hofrichter J (2006) *Biophys J* 91:276–288.
49. Moglich A, Joder K, Kiefhaber T (2006) *Proc Natl Acad Sci USA* 103:12394–12399.
50. Qiu L, Hagen SJ (2004) *J Am Chem Soc* 126:3398–3399.
51. Alm E, Morozov AV, Kortemme T, Baker D (2002) *J Mol Biol* 322:463–476.
52. Kouza M, Li MS, O'Brien EP, Jr, Hu C-K, Thirumalai D (2006) *J Phys Chem* 110:671–676.
53. Naganathan AN, Doshi U, Fung A, Sadqi M, Muñoz V (2006) *Biochemistry* 45:8466–8475.
54. Gilmore R (1981) *Catastrophe Theory for Scientists and Engineers* (Wiley, New York).
55. Wales DJ (2001) *Science* 293:2067–2070.
56. Bogdan TV, Wales DJ (2004) *J Chem Phys* 120:11090–11099.
57. Otzen D, Kristensen O, Proctor M, Oliveberg M (1999) *Biochemistry* 38:6499–6511.
58. Chavez LL, Onuchic JN, Clementi C (2004) *J Am Chem Soc* 126:8426–8432.
59. Wallin S, Chan HS (2006) *J Phys Condens Matter* 18:S307–S328.
60. Lum K, Chandler D, Weeks JD (1999) *J Phys Chem B* 103:4570–4577.
61. Liu Z, Chan HS (2005) *J Mol Biol* 349:872–889.
62. Riddle DS, Grantcharova VP, Santiago JV, Alm E, Ruczinski I, Baker D (1999) *Nat Struct Biol* 6:1016–1024.