Scattered Hammond plots reveal second level of site-specific information in protein folding: ϕ' ($\beta^{\ddagger}$)

Linda Hedberg and Mikael Oliveberg*

Department of Biochemistry, Umeå University, S-901 87 Umeå, Sweden

Edited by Peter G. Wolynes, University of California at San Diego, La Jolla, CA, and approved April 5, 2004 (received for review December 19, 2003)

Site-specific information about structural heterogeneities of the protein-folding transition-state ensemble is commonly derived from the scatter of the Brønsted plot through the individual values of $\phi = \Delta\log k_f/\Delta\log K_{D-N}$. Here, we provide a second level of site-specific detail in the transition-state analysis by demonstrating that the scatter of the Hammond plot is related to heterogeneities in the ϕ-value growth. That is, the extent of transition-state movement ($\Delta\beta^{\ddagger}$) is proportional to the free-energy gradient of the mutational perturbation across the top of the activation barrier, $\phi'(\beta^{\ddagger}) \propto \Delta\log K_{D-N}$. The analysis is applied to the two-state protein L23 where the site-specific free-energy gradients are used to identify the interactions that show the highest degree of consolidation after crossing the barrier top. These interactions are distributed as a shell around the high-ϕ initiation point and denote the side-chain contacts that add criticality to the folding nucleus.

transition state | protein engineering | ϕ-value analysis | Brønsted plot | Hammond behavior

A major challenge in experimental studies of protein folding is to deduce more detailed, high-dimensional information about the diffuse transition-state ensemble to aid the microscopic characterization of folding nucleation and the folding free-energy landscape (1, 2). One strategy has been to see how the structural features of the transition-state ensemble change upon perturbation by denaturants and mutation. A frequently encountered example of such transition-state changes is the Hammond postulate behavior (3). The phenomenon was first reported for barnase and CI2, whose transition-state structures became more native-like as the activation free energy for unfolding was decreased by protein engineering (4, 5). Detection of the transition-state shift was by changes of the kinetic m values, i.e., by small tilts of the v-shaped chevron plots. The Hammond behavior, which has subsequently appeared as a general feature of small proteins (6–16), indicates that the top of the activation barrier is smoothly curved and displays some breadth. Upon mutational perturbation, any slanting of this curvature will cause a movement of the barrier maximum along the folding trajectory (13). From second-order polynomial curvatures of the chevron plots (4), it is further apparent that the barrier tops display an approximately quadratic curvature (13). However, the magnitude of the Hammond shifts for proteins with v-shaped chevron plots is relatively subtle and covers typically <15% of the total folding distance as measured by $\beta^{\ddagger} = m_f/(m_u - m_f)$. Deviation from this minimal behavior is sometimes seen for proteins with pronounced distortions of the v-shaped chevron plot, indicating much larger changes of the transition state position (13, 17). The phenomenon, however, can still be accounted for by low-dimensional movements along the barrier profile, but this time in the form of discrete shifts between consecutive pointed maxima along the folding free-energy profile (7, 14, 18–20). More complex transition-state alterations along secondary order parameters are not normally discerned by the gross parameter $\beta^{\ddagger}$ (6), and seem also to require extensive mutational alterations (7, 21–23) or changes of the sequence connectivity through circular permutation (24–27).

In this study, we present a formalism that quantitatively relates mutational perturbations to the extent of Hammond shifts in protein folding, revealing another aspect of the structural events taking place in the transition-state ensemble: the site-specific growth of contact free energy along the experimental progress coordinate β. The analysis is applied to the ribosomal protein L23 (PDB ID code 1N88; ref. 28), which has not previously been subjected to folding studies. L23 shows v-shaped chevron plots and a classical nucleation-condensation distribution of fractional ϕ values in the transition state. The mutant-induced m-value changes show further that L23 describes moderate Hammond shifts, but no significant plasticity of the denatured or native ground states. The critical feature of the L23 Hammond plot is, however, that the individual m values are scattered beyond the experimental errors, allowing a detailed description of the contact growth within the critical nucleus as the protein traverses the barrier top.

Theory

General Features of Two-State Folding. In two-state folding, the equilibrium constant ($K_{D-N} = [D]/[N]$) has been shown to equal the ratio of the unfolding and refolding rate constants (29), $K_{D-N} = k_u/k_f$, where

$$\log K_{D\text{-}N} = \log K_{D\text{-}N}{}^{H_2O} + m_{D\text{-}N}[\mathrm{GdmCl}]\,, \qquad [1]$$

$$\log k_u = \log k_u{}^{H_2O} + m_u[\mathrm{GdmCl}], \text{ and} \qquad [2]$$

$$\log k_f = \log k_f{}^{H_2O} + m_f[\mathrm{GdmCl}], \qquad [3]$$

according to the empirical observation of linear free-energy changes after the addition of chemical denaturants like GdmCl or urea (30). The parameters are derived experimentally from chevron plots; i.e., plots of log k_f and log k_u versus [GdmCl], fitting the equation

$$\log k_{obs}\ \log(10^{\log k_f^{H_2O} + m_f[\mathrm{GdmCl}]}10^{\log k_u^{H_2O} + m_u[\mathrm{GdmCl}]}) \qquad [4]$$

The slopes $m_u - m_f = m_{D-N}$ are commonly taken as measures of the change in solvent exposure in the activation process of unfolding and refolding, respectively. The m values are used to position the folding transition-state ensemble on the experimental progress coordinate β according to refs. 29 and 31:

$$\beta^{\ddagger} = m_f/m_{D\text{-}N} = 1 - m_u/m_{D\text{-}N}\,. \qquad [5]$$

Additional, site-specific information about the contacts present in the transition-state structure is obtained from the ϕ values (30)

$$\phi = \Delta\Delta G^{\ddagger}/\Delta\Delta G_{D\text{-}N} = \Delta\log k_f/(\Delta\log k_u - \Delta\log k_f) \qquad [6]$$

that measure the effect of point mutation on the transition-state stability ($\Delta\Delta G^{\ddagger}$) relative to that of the native state ($\Delta\Delta G_{D-N}$). A ϕ value of 1 indicates that the structure around the mutated side chain has grown fully native-like at the position of

This paper was submitted directly (Track II) to the PNAS office.

*To whom correspondence should be addressed. E-mail: mikael.oliveberg@chem.umu.se.

www.pnas.org/cgi/doi/10.1073/pnas.0308497101

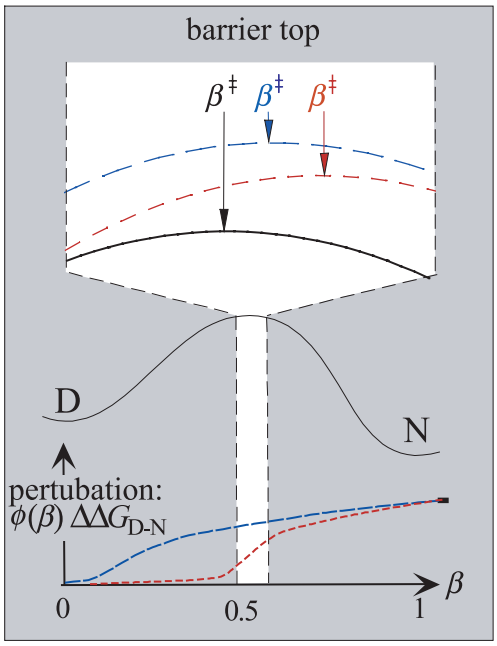

Fig. 1. Schematic illustration of transition-state shift on a smooth folding free-energy profile, including an expanded view of the barrier top. After mutational perturbation, the barrier profile is offset by $\Delta\Delta G(\beta) = \phi(\beta)\Delta\Delta G_{D-N}$ (Eq. **7**), describing the loss of interaction free energy along the experimental coordinate β. The resulting extent of transition-state shift is then determined by the slope of the perturbation across the barrier top (Eqs. **7–11**). Exemplified are two mutations with different values of ϕ and different gradients of $\Delta\Delta G(\beta)$ across the narrow interval of the transition-state shift (blue and red). For simplicity, the mutations are assumed to have the same $\Delta\Delta G_{D-N}$. The red mutation causes larger shifts of $\beta^{\ddagger}$ than the blue because its relative effect of the early and late part of the barrier top is larger. Mutations with constant values of ϕ across the barrier top do not produce transition-state shifts.

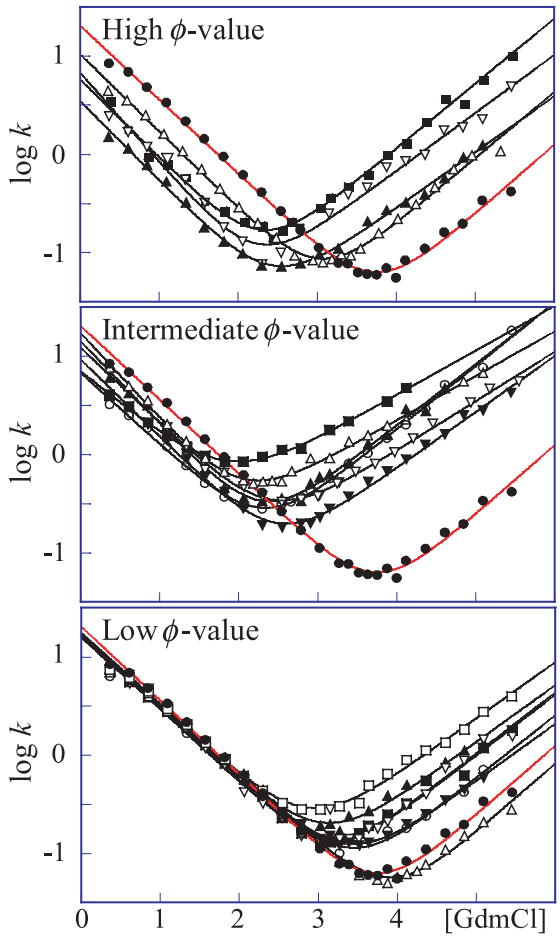

Fig. 2. Chevron plots for wild-type (●) and mutant L23, grouped according to their ϕ values. Units for the rate constants are in s^{-1}. (*Top*) Mutants with $\phi >$ 0.25. V30A (■), V81A (▲), V43A (▽), and I39V (△). (*Middle*) Mutants with $0.25 > \phi > 0.10$. Y26A (■), V49A (▲), V51A (▽), V54A (△), I80A (▼), and V83A (○). (*Bottom*) Mutants with $\phi < 0.10$. I8V (■), V12A (▲), L13A (▽), Y18A (△), F21A (▼), L57A (○), and I89A (□). In Figs. 3 and 4, mutants with high, intermediate, and low ϕ values are red, orange, and yellow, respectively. In Table 1, mutants with high and intermediate ϕ values are bold and underlined, respectively.

the transition state ($\beta = \beta^{\ddagger}$), whereas a ϕ value of 0 indicates that it is still disordered.

Basic Rationale for Hammond Postulate Behavior in Protein Folding. The relation between mutational perturbation and transition-state movements; i.e., changes of $\beta^{\ddagger}$, was deduced as follows. We have previously demonstrated, by monitoring the effect of individual mutations along the broad barrier region of U1A, that the stability loss on mutation increases progressively with β (15). That is, mutations have generally larger energetic effect on the late part of the reaction profile where the interactions are better consolidated. A general expression for this mutant-induced perturbation is given by

$$\Delta\Delta G\,(\beta) = \phi\,(\beta)\Delta\Delta G_{D\text{-}N}\,, \qquad [7]$$

where $\Delta\Delta G(\beta)$ is the stability loss as a function of β and $\phi(\beta)$ describes the ϕ-value growth along β (Fig. 1). Because $\Delta\Delta G(\beta)$ increases along the experimental progress coordinate, the free-energy profile is shifted up more near the native state and, in combination with any width of the barrier top, this biased perturbation will cause the transition state to shift closer to the native state. For quadratic curvatures of the barrier top, as is suggested experimentally from second-order polynomial dependence of $\log k_u$ on [GdmCl] (13), the folding free-energy profile near the transition state may be approximated with the function

$$\Delta G(\beta) = \Delta G(\beta^{\ddagger}) + c(\beta - \beta^{\ddagger})^2\,, \qquad [8]$$

where $c < 0$. The transition-state shift following the mutational perturbation $\Delta G^{mut}(\beta) = \Delta G(\beta) + \phi(\beta)\Delta\Delta G_{D-N}$ is then given by $0 = \Delta G'(\beta) + \phi'(\beta)\Delta\Delta G_{D-N} \Rightarrow$

$$\Delta\beta^{\ddagger} = (\beta^{\ddagger\,\mathrm{mutant}} - \beta^{\ddagger wt}) = -\phi'(\beta^{\ddagger wt})\Delta\Delta G_{D\text{-}N}/2c, \qquad [9]$$

and from Eq. **5**

$$\Delta m_u \propto \phi'(\beta^{\ddagger wt})/\Delta\Delta G_{D\text{-}N}\,, \qquad [10]$$

assuming that the perturbation $\phi(\beta)\Delta\Delta G_{D-N}$ is approximately linear in the narrow interval $\Delta\beta^{\ddagger}$. The corresponding barrier height at $\beta^{\ddagger mut}$ is

$$\begin{aligned}\Delta G^{mut}(\beta^{\ddagger mut}) &= \Delta G(\beta^{\ddagger mut}) + \phi(\beta^{\ddagger mut})\Delta\Delta G_{D\text{-}N} \\ &= \Delta G(\beta^{\ddagger wt}) + \phi(\beta^{\ddagger wt})\Delta\Delta G_{D\text{-}N} - (\phi'(\beta^{\ddagger wt})\Delta\Delta G_{D\text{-}N})^2/4c\end{aligned} \qquad [11]$$

In other words, the extent of transition-state movement (Δm_u) is proportional to the gradient of the mutational perturbation across the free-energy maximum. Hammond shifts are thus not expected for mutations with constant values of ϕ across the barrier top, because these mutations will just contribute to offset

Table 1. Parameters from kinetic measurements

Mutant	$\log k_f^{H_2O}$	$\log k_u^{H_2O}$	Midpoint (M)	m_f	m_u	$m_{D\text{-}N}$ ($mol^{-1}\cdot M^{-1}$)	$\Delta G_{D\text{-}N}$(kcal/mol^{-1})	$\Delta\Delta G_{D\text{-}N}$ (kcal/mol^{-1})	$\phi^{1.5M}$	$\beta^{\ddagger}$	Δm_u	$\Delta m_u/\Delta \log K_{D\text{-}N}$
WT	1.31 ± 0.03	−4.13 ± 0.02	3.73 ± 0.05	−0.75 ± 0.02	0.70 ± 0.04	1.46 ± 0.05	7.39 ± 0.28			0.52		
18V	1.22 ± 0.04	−3.10 ± 0.16	3.17 ± 0.07	−0.74 ± 0.03	0.62 ± 0.03	1.36 ± 0.04	5.87 ± 0.22	1.52 ± 0.36	0.08	0.55	0.09	**0.11**
V12A	1.25 ± 0.04	−2.76 ± 0.16	3.06 ± 0.07	−0.73 ± 0.03	0.58 ± 0.04	1.31 ± 0.04	5.45 ± 0.22	1.94 ± 0.36	0.02	0.56	0.12	**0.14**
L13A	1.21 ± 0.06	−3.10 ± 0.21	3.15 ± 0.09	−0.74 ± 0.04	0.62 ± 0.05	1.37 ± 0.06	5.87 ± 0.29	1.52 ± 0.41	0.08	0.54	0.08	**0.10**
Y18A	1.20 ± 0.03	−4.27 ± 0.23	3.91 ± 0.05	−0.70 ± 0.02	0.70 ± 0.05	1.40 ± 0.05	7.45 ± 0.31	−0.06 ± 0.42		0.50	0.01	−0.02
F21A	1.24 ± 0.03	−3.17 ± 0.14	3.39 ± 0.05	−0.72 ± 0.01	0.58 ± 0.03	1.30 ± 0.04	6.00 ± 0.20	1.39 ± 0.34	0.03	0.55	0.12	**0.26**
Y26A	0.84 ± 0.02	−1.20 ± 0.06	1.85 ± 0.05	−0.65 ± 0.03	0.45 ± 0.02	1.10 ± 0.03	2.78 ± 0.09	4.61 ± 0.29	0.11	0.59	**0.25**	**0.10**
V30A	0.76 ± 0.05	−2.58 ± 0.09	2.30 ± 0.06	−0.80 ± 0.04	0.66 ± 0.02	1.46 ± 0.05	4.55 ± 0.14	2.84 ± 0.31	**0.29**	0.55	0.04	0.02
I39V	1.02 ± 0.04	−3.40 ± 0.12	3.03 ± 0.05	−0.79 ± 0.02	0.67 ± 0.03	1.46 ± 0.04	6.00 ± 0.18	1.39 ± 0.33	**0.33**	0.54	0.03	0.03
V43A	0.83 ± 0.09	−2.53 ± 0.17	2.24 ± 0.10	−0.91 ± 0.07	0.59 ± 0.04	1.50 ± 0.08	4.58 ± 0.26	2.82 ± 0.38	**0.34**	0.61	0.11	0.06
V49A	1.15 ± 0.05	−2.20 ± 0.11	2.29 ± 0.06	−0.84 ± 0.04	0.62 ± 0.03	1.46 ± 0.05	4.56 ± 0.16	2.83 ± 0.32	0.14	0.57	0.08	0.04
V51A	1.09 ± 0.08	−1.97 ± 0.06	2.41 ± 0.06	−0.76 ± 0.05	0.50 ± 0.01	1.27 ± 0.05	4.15 ± 0.14	3.24 ± 0.31	0.12	0.60	0.20	**0.11**
V54A	1.24 ± 0.03	−1.46 ± 0.06	2.03 ± 0.05	−0.88 ± 0.03	0.45 ± 0.01	1.33 ± 0.04	3.66 ± 0.09	3.73 ± 0.29	0.10	0.66	**0.25**	**0.11**
L57A	1.19 ± 0.04	−3.42 ± 0.31	3.42 ± 0.09	−0.71 ± 0.03	0.64 ± 0.07	1.35 ± 0.07	6.28 ± 0.43	1.11 ± 0.51	0.08	0.53	0.07	**0.15**
I80A	0.97 ± 0.03	−2.42 ± 0.07	2.49 ± 0.05	−0.79 ± 0.03	0.57 ± 0.02	1.36 ± 0.03	4.60 ± 0.11	2.79 ± 0.30	0.21	0.58	0.13	**0.08**
V81A	0.54 ± 0.04	−2.82 ± 0.09	2.44 ± 0.05	−0.81 ± 0.03	0.57 ± 0.02	1.38 ± 0.04	4.57 ± 0.13	2.82 ± 0.31	**0.44**	0.59	0.14	0.08
V83A	0.83 ± 0.03	−2.32 ± 0.06	2.30 ± 0.04	−0.73 ± 0.03	0.65 ± 0.01	1.37 ± 0.03	4.29 ± 0.09	3.10 ± 0.29	0.21	0.53	0.06	0.03
I89A	1.22 ± 0.03	−2.49 ± 0.10	2.88 ± 0.05	−0.71 ± 0.02	0.57 ± 0.02	1.29 ± 0.03	5.04 ± 0.15	2.35 ± 0.32	0.02	0.55	0.13	**0.11**

$\phi^{1.5M}$ calculated from refolding and unfolding data at 1.5 M GdmCl (Eq. **6**) High and intermediate values are bold and underlined, respectively.

the barrier without affecting its tilt. $\beta^{\ddagger}$ shifts arise only from mutation of interactions that undergo changes upon traversing the barrier top, i.e., interactions that are in the process of being formed (5). Preformed or fully unfolded parts of the protein are just carried over silently. Notably, it can be shown that the proportionality between Δm_u and $\phi'(\beta^{\ddagger\,wt})\Delta\Delta G_{D-N}$ is qualitatively the same whether the barrier top is not perfectly smooth as assumed in Eq. **8**, but riddled with fine structure. The values of Δm_u would then tend to change discretely with $\Delta\Delta G_{D-N}$ as the transition state shifts from one local maximum to the next. However, the approximately smooth distribution of the experimentally observed m_u values provides no evidence for the existence of such fine structure (compare Fig. 5). To this end, it may further be argued that, for the small β-windows studied here, the fine structure should be smoothened out by conformational heterogeneities of the transition-state ensemble. Another notable feature of the simplistic formalism in Fig. 1 is that it captures the experimental finding that, in protein folding, the Hammond shift ($\Delta\beta^{\ddagger}$ or Δm_u) is linearly related to $\Delta\Delta G_{D-N}$ (13) and not to changes in the barrier height (Eqs. **8–11**).

Experiment

Materials. Mutagenesis was performed by standard procedures using the QuikChange kit (Stratagene). The protein was expressed in *Escherichia coli* (BL21 DE3) and purified with cation exchange (SP-Sepharose, Amersham Pharmacia) and gel filtration (S100, Amersham Pharmacia). The stopped-flow measurements were carried out on a SX-17MV instrument (Applied Photophysics, Surrey, U.K.) at 25°. Buffer was 50 mM Mes at pH 6.3 (Sigma) and 0.23 M Na_2SO_4 (Merck). The guanidinium chloride (GdmCl) was ultra pure from ICN Biomedicals.

Na_2SO_4 Allows Analysis of Marginally Stable Proteins. Because the stability of wild-type L23 is relatively low, the experiments were performed under mildly stabilizing conditions in the presence of 0.23 M Na_2SO_4 to assure accurate determination of the refolding rate constant and the refolding m value (Eqs. **2** and **3**). Under these conditions, L23 displays a transition midpoint of 3.73 M GdmCl and $\Delta G_{D\text{-}N}$ = 7.4 kcal/mol, corresponding to a stabilizing power of Na_2SO_4 of ≈3 (kcal/mol) M. In the absence of Na_2SO_4, midpoint = 2.77 M and ΔG_{D-N} = 4.6 kcal/mol.

A side effect of subjecting proteins to Na_2SO_4, or other stabilizing agents, is that the increased contact strength may lead to premature collapse of the coil, revealed by rollovers in the refolding limb of the chevron plot (32, 33). Although such rollovers may be discerned >0.5 M Na_2SO_4, the effect is negligible at 0.23 M. As a control, the ϕ values for several mutants were determined in both 0 and 0.23 M Na_2SO_4 and found to be the same within the experimental errors. Moreover, the exclusion of chevron data <1 M GdmCl does not significantly affect the m value changes upon mutation.

V-Shaped Chevrons and Diffuse Nucleus Suggest "Classical" Two-State Behavior. L23 consists of a central four-stranded β-sheet that is clamped by three helices forming a saddle-shaped hydrophobic core (28). A notable feature of this core is that the aromatic side chains are polarized toward the N-terminal side, whereas the C-terminal side consists mainly of aliphatic moieties. For the analysis, 17 conservative mutations were performed throughout the interior of the L23 structure, all of which showed approximately v-shaped chevron plots with overall similar, but not identical, m values (Fig. 2). By this criterion, L23 classifies as a classical two-state folder (34). Even so, the slightly different tilts of the chevron plots, i.e., changes in m_f and m_u (Table 1), will produce a slight GdmCl dependence of the ϕ value according to Eq. **6**. The phenomenon is apparent to various extent in all two-state proteins and seems to reflect structural reorganizations of the transition state upon mutation (35). To minimize the consequences of these m-value changes, we have estimated the ϕ values near the transition midpoint at 1.5 M GdmCl, where the refolding rate constant (k_f) can be measured directly and the extrapolation of unfolding data (k_u) is minimal. All L23 mutations display fractional ϕ values, ranging from a highest value of 0.44 for V81A to nearly zero for V12A and I89A (Table 1). This diffuse nature of the L23 transition state is further apparent from the Brønsted plot (i.e., $\Delta\log k_u$ versus $\Delta\log K_{D-N}$), which yields $<\phi> = 0.18$ (Fig. 3). When calculated at other GdmCl concentrations, the internal graduation of the ϕ values remains largely the same, showing that the deviation from "ideal" two-state behavior of L23 is too small to have any significant bearing on the structural analysis of the transition-state ensemble. The stability loss of the mutants in Table 1 varies from 1.1 kcal/mol (L57A) to >4 kcal/mol (Y26A),

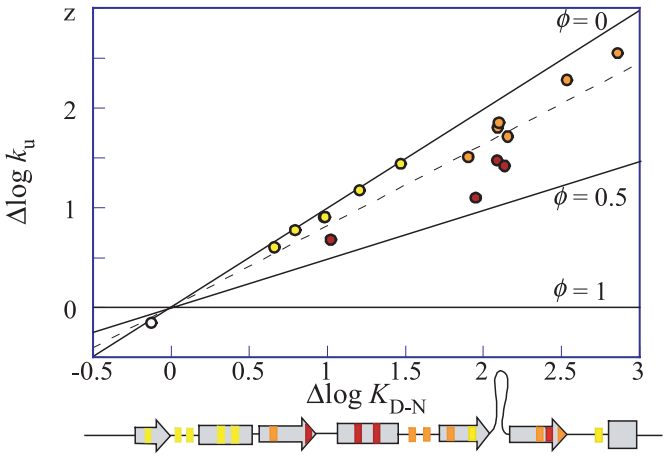

Fig. 3. Brønsted plot, logk_u versus logK_{D-N}, of L23 at 1.5 M GdmCl. The ϕ values for the individual mutations are calculated from Eq. **6** and are colored as in Fig. 4. The dotted line is $<\phi> = 0.18$. Below is the mutant position in the L23 primary sequence.

with the exception of the excluded mutation Y18A that shows too slight a stability change to allow accurate determination of its ϕ-value. Taken together, these straightforward experimental characteristics make L23 an ideal model system for tests of the formalism in Fig. 1.

ϕ Value Analysis of the L23 Transition State. The positions of the mutated residues in the L23 structure are shown in Fig. 4, where the individual ϕ values are coded yellow ($0<\phi<0.10$), orange ($0.10<\phi<0.25$), and red ($\phi>0.25$). The thresholds have been chosen for clarity. Apparently, the L23 ϕ values describe an approximately radial graduation from the center of the hydrophobic core, which is in good agreement with the nucleation condensation mechanism for two-state folding (36, 37). The highest ϕ values (red) are observed for the triad Val-81 ($\phi = 0.44$), Ile-39 ($\phi = 0.33$) and Val-43 ($\phi = 0.34$), linking the central strand 4 and helix 2 by contacts through the aliphatic side of the core, flanked by Val-30 ($\phi = 0.29$) in strand 3 closing the loop to helix 2. In direct connection to this high-ϕ cluster is a contiguous group of residues with intermediate ϕ values (orange), closing up the central part of the β-sheet by connecting strands 2, 4, and 3; i.e., Tyr-26 ($\phi = 0.11$), Val-49 ($\phi = 0.14$), Val-51 ($\phi = 0.12$), Val-54 ($\phi = 0.10$), Ile-80 ($\phi = 0.21$), and Val-83 ($\phi = 0.21$). The lowest ϕ values (yellow) are found for the more peripheral residues Ile-8 ($\phi = 0.08$), Val-12 ($\phi = 0.02$), Leu-13 ($\phi = 0.08$), Phe-21 ($\phi = 0.03$), and Ile-89 ($\phi = 0.02$), anchoring the N and C termini, and Leu-57 ($\phi = 0.08$) near the disordered loop.

Matched *m* Value Changes Indicative of Transition-State Movements. Transition-state movements in protein folding are commonly inferred from changes in $\beta^{\ddagger}$ (Eq. **5**), following stability perturbations by either mutation or addition of denaturant (5, 13). However, the interpretation of these $\beta^{\ddagger}$ changes is not always straightforward because they may also result from alterations of the ground states D or N (8, 19). To test for such ground-state alterations in the case of L23, we first replaced $\beta^{\ddagger}$ with m_u that measures directly the "distance" between the native starting material and the transition state for unfolding (Eq. **2**). The plot of m_u versus change in unfolding activation free energy ($\propto \Delta\log k_u$) for the 17 mutations in Table 1 shows a significant linear correlation ($R = 0.82$), which is in apparent agreement with the Hammond postulate (Fig. 5). Because the correlation is largely retained in a plot of $\beta^{\ddagger}$ versus $\Delta\log K_{D-N}{}^{mut}$ ($R = 0.74$), we

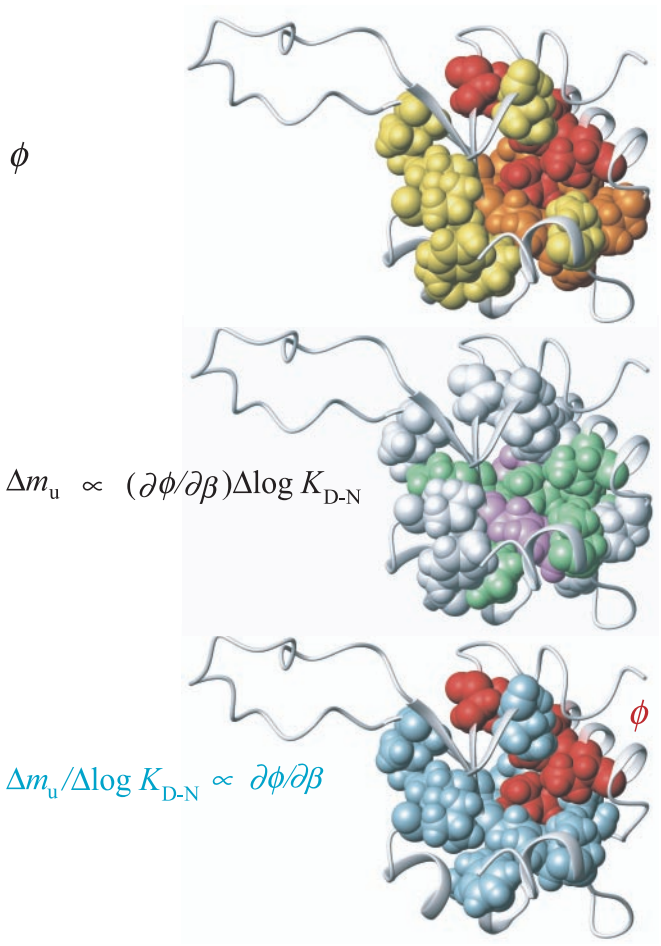

Fig. 4. L23 structures showing spatial distribution of ϕ, Δm_u, and $\Delta m_u/\Delta\Delta G_{D-N}$. (*Top*) The highest values of ϕ (red) are found in the aliphatic side of the L23 core, which is consistent with a prenucleus of long-range contacts between Val-81 in the central strand 4, and residues in the helix-2 region. (*Middle*) The highest values of Δm_u (magenta) are found for the sheet residues on either side of Val-81. Depicted are the residues that show the highest degree of free-energy gain after passing the barrier top. The intermediate values of Δm_u (green) are found in the central region of the L23 core, whereas the lowest values of Δm_u (white) are found at more peripheral positions, including the interfaces to the N and C termini. (*Bottom*) In the diffuse transition-state ensemble of L23, the highest ϕ-value growth, $\Delta m_u/\Delta\log K_{D-N} \propto \phi'(\beta^{\ddagger})$, is observed for residues just outside the initiation site (red). Blue residues are those with $\Delta m_u//\Delta\log K_{D-N}$ larger than within the high-ϕ cluster (red).

conclude that contributions from denatured-state plasticity is after all very small with L23 (Fig. 5). The increased values of m_f are matched by a decrease in m_u, as would be expected for transition-state movements (Table 1). As an alternative test, the average ϕ value ($<\phi>$) derived from the slope of the Brønsted plot was used as a measure of transition-state position and plotted against $\Delta\log K_{D-N}{}^{GdmCl} = <m_{D-N}>[GdmCl]$, where $<m_{D-N}>$ is the average value of m_{D-N} in Table 1. The result shows a linear correlation with a slope that is within the experimental error of that for $\beta^{\ddagger}$ versus $\Delta\log K_{D-N}{}^{mut}$ (Fig. 5). It is thus apparent that $<\phi>$ and $\beta^{\ddagger}$ produce coherent Hammond shifts although their overall values for the transition-state placement are very different at ≈0.18 and ≈0.5, respectively.

Scatter in Hammond Plot Exceeds Experimental Errors. In view of the smooth transition-state shifts observed upon addition of denaturant (13), the plot of m_u versus $\Delta\log k_u$ reveals considerable

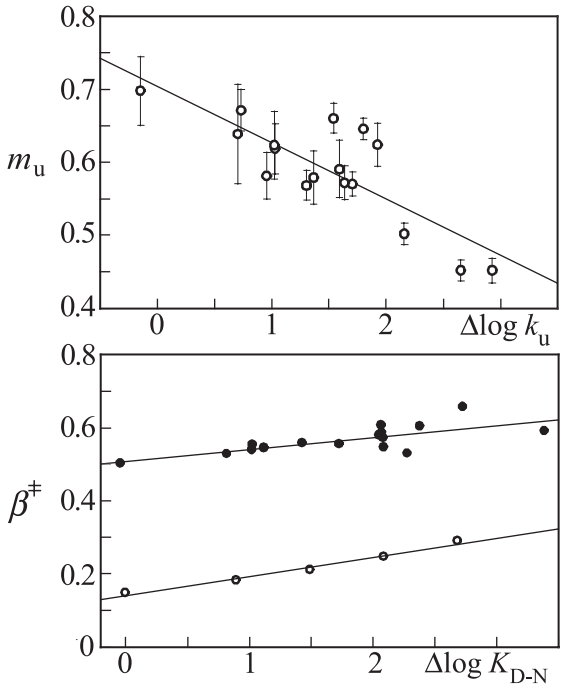

Fig. 5. *m* values change after mutational perturbation of L23. (*Upper*). Hammond plot (m_u versus $\Delta\log k_u$) showing that the scatter of the m_u values is larger than the experimental errors. The fitted line is $m_u = 0.7–0.077(\pm0.01)\Delta\log k_u$, $R = 0.82$. (*Lower*) Comparison of the Hammond plots $\beta^{\ddagger}$ versus $\Delta\log K_{D-N}$ from data in Table 1 (●) and the Brønsted slopes ($<\phi>$ in Fig. 3) versus $\Delta\log K_{D-N} = <m_{D-N}>$[GdmCl] (○). $\beta^{\ddagger} = 0.5 + 0.033(\pm 0.01)\Delta\log K_{D-N}$ ($R = 0.74$) and $<\phi> = 0.14 + 0.053(\pm 0.004)\Delta\log K_{D-N}$ ($R = 0.99$). Although the absolute values for the transition-state position are different, their dependence on $\Delta\log K_{D-N}$ is overall similar, providing additional evidence for transition state movements in L23.

scatter (Fig. 5). Notably, this scatter does not arise from experimental errors but shows mutant-specific deviations from the mean kinetic pattern. Compare, for example, I80A and V83A in Fig. 2 that have similar $\Delta\log k_u$ but distinctly different m_u. It can further be noted that the outliers in the L23 Hammond plot are not correlated with their extent of interaction in the transition-state ensemble, i.e., anomalously small m_u changes are observed for both high and low ϕ values (Table 1). Scattered Hammond plots are not unique for L23 but have also been observed for barnase and CI2, where the susceptibility to undergo transition-state movements was found to be different in different parts of the structure (4, 5). Even so, these deviations from uniform Hammond behavior cannot *a priori* be attributed to transition state effects alone. One possibility is that they arise from partial unfolding of the native protein. Such ground-state fraying has been presented as an alternative explanation for kinked $\log k_u$ limbs of certain S6 mutations, based on their location at the late-forming interface between the N and C termini (19). However, this interpretation seems unlikely in the case of L23 because the largest m_u changes involve core mutations where native-state fraying seems incompatible with microscopic reversibility (Fig. 4). Moreover, the mutations at the seemingly late folding N- and C-terminal regions of L23 show the smallest m_u changes (Fig. 4). Another possibility is that the Hammond shifts describe redistributions between parallel folding pathways that extend beyond the minimal two-dimensional barrier model (4, 6). In accord with the prediction in Eq. **10**, however, we opt for a more rational explanation: the scattered Hammond plots in protein folding emerge as an intrinsic consequence of nonuniform values of ϕ' ($\beta^{\ddagger\,wt}$) $\Delta\Delta G_{D-N}$ (Fig. 1).

Discussion

The Diffuse Transition-State Ensemble. The ϕ-value data of L23 yields an overall diffuse impression of the transition state with a graduated distribution of fractional ϕ values that encompasses the majority of the native contacts (37). This delocalized nucleation pattern is a typical feature of two-state proteins (34, 36, 37) and is generally taken to indicate that the folding transition state is an ensemble of states whose average topology is native-like (38, 39). Consistently, the logarithm of the refolding rate constant across different two-state proteins is found to correlate with the topological parameter "contact order" of the native state (40), with some intrinsic modulation by protein stability and chain length (40–44). The native-like appearance of the transition-state ensemble suggests that the various parts of a protein fold with similar probability (38). At the extreme, such equalization would lead to a situation where all folding and unfolding trajectories are equally favorable, producing ideal Brønsted plots with uniform ϕ values and maximal cooperativity; local unfolding of the native structure is entropically prevented because all native interactions are fractionally present in the transition state (25). In practice, however, the graduated ϕ-value distribution indicates often a dominant initiation point that can be qualitatively predicted by simplistic topology-based folding models (45–48). Simulations suggest further that this initiation point, although it looks diffuse by ϕ values, could still be consistent with a subset of fully established contacts that are shared by the majority of species in the transition-state ensemble (49). With L23, the initiation point is the docking of the helix 2 and strand 4 regions by long-range contacts through the aliphatic side of the core (Fig. 4). Concomitant with this long-range hydrophobic association is probably the local trapping of hydrogen bonds in helix 2 that are "flickeringly" present in the denatured ensemble (37, 50). Structural rationalization of the subsequent layers of contacts becomes increasingly degenerate as the configurational dimensionality grows rapidly with decreasing values of ϕ (39). Even so, additional information about the folding nucleation events can be deduced from the m_u value changes in Fig. 5.

Variable Hammond Response Linked to Heterogeneous ϕ Value Growth. In contrast to the low-dimensional case of covalent bonds, it is evident that the Hammond behavior in protein folding is governed by multiple order parameters: the global progress coordinate β by which the transition-state shift is measured, and a large set of site-specific coordinates $\Delta\Delta G(\beta) = \phi(\beta)\Delta\Delta G_{D-N}$ that determine the energetic perturbations upon mutation. It follows that the individual data points in the Hammond plot (Δm_u versus $\Delta\Delta G_{D-N}$), which are derived from site-specific free-energy perturbations, reflect local properties $\propto\phi'(\beta^{\ddagger})$. The common Hammond plot, based on a superposition of data from structurally dispersed mutations, is therefore expected to contain scatter unless the local values of $\phi'(\beta^{\ddagger})$ are the same throughout the transition-state structure. However, such uniform growth of the local progress coordinates seems incompatible with the way real proteins fold. The ϕ values are different throughout the protein structure and are bound to show nonlinear dependencies on β. Thus, the individual values of $\phi'(\beta^{\ddagger})$ are also likely to be different at any given point along the reaction profile, which is consistent with experimental data on U1A (15).

On this basis, we conclude that the scatter of the Hammond plot (Fig. 5) is related to site-specific variations in $\phi'(\beta^{\ddagger})$, similar to the way the scatter of the Brønsted plot is related to the site-specific values of $\phi(\beta^{\ddagger})$ (Fig. 3).

An Additional Facet of the Folding Nucleus: The Critical Contact Layer. To examine the mechanistic implications of the Hammond data, we have tentatively defined $\Delta m_u \propto \phi'(\beta^{\ddagger})$ as the site-specific

free-energy gradient, and the normalized ditto $\Delta m_u/\Delta\Delta G_{D-N}\phi'(\beta^{\ddagger})$ as the site-specific ϕ value gradient. Absolute estimates of these parameters require precise knowledge about the barrier curvature, i.e., the value of c in Eq. **8**. Accordingly, Δm_u can be seen as a directional parameter indicating in which part of the transition-state structure the growth of interaction free energy is most downhill, i.e., the interactions that show the largest degree of consolidation after traversing the barrier top. $\Delta m_u/\Delta\Delta G_{D-N}$, on the other hand, represents a local order parameter indicating heterogeneities in the structural growth as defined by Eq. **6**. The graduation of these parameters throughout the L23 structure is shown in Fig. 4, where the individual values are coded: magenta ($\Delta m_u > 0.21$), green ($0.1 < \Delta m_u < 0.20$), white ($\Delta m_u < 0.1$), and pale blue ($\Delta m_u/\Delta\Delta G_{D-N} > 0.079$). The result shows that the distribution of Δm_u values is different from that of the ϕ values. Whereas the center of the folding nucleus is in the aliphatic side of the core, the high values of Δm_u are, on the whole, shifted toward the aromatic side (Fig. 4). The implication of this spatial separation of the Δm_u and ϕ value distributions is that, at the barrier top, the local free-energy gradients are steeper at the interface to the folding nucleus than in its center. Looking at a more detailed level the pattern emerges even clearer. The two highest Δm_u values (> 0.21) are found for the sheet mutations Y26A ($\phi = 0.11$, strand 2) and V54A ($\phi = 0.10$, strand 3), flanking the initiation site V81 ($\phi = 0.44$) in the central strand 4, whereas the next layer of residues with intermediate Δm_u values covers the aromatic part of the core. The lowest Δm_u values are observed at the interface to the N and C termini, and at positions peripheral to the sheet. Notably, this low-Δm_u set includes also the two residues in the high-ϕ cluster facing the C terminus, i.e., V30A ($\phi = 0.29$) and I39V ($\phi = 0.33$), emphasizing the polarization of high Δm_u values toward side chains closing up the β-sheet.

The accompanying values of $\Delta m_u/\Delta\Delta G_{D-N}$ yield a very similar picture. The ϕ value growth is higher for the residues distributed around the initiation point than in its center (Fig. 4), in apparent analogy with the "hydrophobic cloud" wetting the central nucleus in the capillarity description of protein folding (51).

Accordingly, the results allow us to make a crude distinction between a primary set of interactions that provides the gross topology for nucleation; i.e., those that form on the uphill side of the free-energy barrier, and a secondary set that brings the process downhill. In somewhat simplified terms, the latter, deterministic, interactions are those that make the embryonic nucleus critical: the interactions that undergo the highest free-energy change after passing the barrier top. For comparison, the critical nucleus of water condensation is solely defined by size, because the individual molecules within the emerging droplet are indistinguishable. In protein folding, the interactions that add criticality to the nucleus have identity: the critical contact layer.

We thank Håkan Wennerström and Anders Öhman for stimulating discussions and the Swedish Research Council for financial support.

1. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995) *Proteins* **21,** 167–195.
2. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J. & Socci, N. D. (2000) *Adv. Protein Chem.* **53,** 87–152.
3. Hammond, G. S. (1955) *J. Am. Chem. Soc.* **77,** 334–338.
4. Matouschek, A. & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7814–7818.
5. Matouschek, A., Otzen, D. E., Itzhaki, L. S., Jackson, S. E. & Fersht, A. R. (1995) *Biochemistry* **34,** 13656–13662.
6. Matthews, J. M. & Fersht, A. R. (1995) *Biochemistry* **34,** 6805–6814.
7. Crane, J. C., Koepf, E. K., Kelly, J. W. & Gruebele, M. (2000) *J. Mol. Biol.* **298,** 283–292.
8. Sanchez, I. E. & Kiefhaber, T. (2003) *J. Mol. Biol.* **327,** 867–884.
9. Sanchez, I. E. & Kiefhaber, T. (2003) *Biophys. Chem.* **100,** 397–407.
10. Yiu, C. P., Mateu, M. G. & Fersht, A. R. (2000) *Chembiochem.* **1,** 49–55.
11. Isaacson, R. L., Weeds, A. G. & Fersht, A. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11247–11252.
12. Dalby, P. A., Oliveberg, M. & Fersht, A. R. (1998) *Biochemistry* **37,** 4674–4679.
13. Silow, M. & Oliveberg, M. (1997) *Biochemistry* **36,** 7633–7637.
14. Otzen, D. E., Kristensen, O., Proctor, M. & Oliveberg, M. (1999) *Biochemistry* **38,** 6499–6511.
15. Ternstrom, T., Mayor, U., Akke, M. & Oliveberg, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14854–14859.
16. Portman, J. J., Takada, S. & Wolynes, P. G. (1998) *Phys. Rev. Lett.* **81,** 5237–5240.
17. Oliveberg, M. (1998) *Acc. Chem. Res.* **31,** 765–772.
18. Oliveberg, M., Tan, Y. J., Silow, M. & Fersht, A. R. (1998) *J. Mol. Biol.* **277,** 933–943.
19. Otzen, D. E. & Oliveberg, M. (2002) *J. Mol. Biol.* **317,** 639–653.
20. Sanchez, I. E. & Kiefhaber, T. (2003) *J. Mol. Biol.* **325,** 367–376.
21. Nauli, S., Kuhlman, B. & Baker, D. (2001) *Nat. Struct. Biol.* **8,** 602–605.
22. Daggett, V. & Fersht, A. R. (2003) *Trends Biochem. Sci.* **28,** 18–25.
23. Gianni, S., Guydosh, N. R., Khan, F., Caldas, T. D., Mayor, U., White, G. W., DeMarco, M. L., Daggett, V. & Fersht, A. R. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 13286–13291.
24. Viguera, A. R., Blanco, F. J. & Serrano, L. (1995) *J. Mol. Biol.* **247,** 670–681.
25. Lindberg, M., Tangrot, J. & Oliveberg, M. (2002) *Nat. Struct. Biol.* **9,** 818–822.
26. Li, L. & Shakhnovich, E. I. (2001) *J. Mol. Biol.* **306,** 121–132.
27. Weikl, T. R. & Dill, K. A. (2003) *J. Mol. Biol.* **332,** 953–963.
28. Ohman, A., Rak, A., Dontsova, M., Garber, M. B. & Hard, T. (2003) *J. Biomol. NMR* **26,** 131–137.
29. Jackson, S. E. & Fersht, A. R. (1991) *Biochemistry* **30,** 10428–10435.
30. Fersht, A. R. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
31. Tanford, C. (1970) *Adv. Protein Chem.* **24,** 1–95.
32. Otzen, D. E. & Oliveberg, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11746–11751.
33. Silow, M. & Oliveberg, M. (2003) *J. Mol. Biol.* **326,** 263–271.
34. Jackson, S. E. (1998) *Folding Des.* **3,** R81–R91.
35. Oliveberg, M. (2001) *Curr. Opin. Struct. Biol.* **11,** 94–100.
36. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1994) *Biochemistry* **33,** 10026–10036.
37. Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10869–10873.
38. Plotkin, S. S. & Onuchic, J. N. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6509–6514.
39. Vendruscolo, M., Paci, E., Dobson, C. M. & Karplus, M. (2001) *Nature* **409,** 641–645.
40. Plaxco, K. W., Simons, K. T. & Baker, D. (1998) *J. Mol. Biol.* **277,** 985–994.
41. Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D. & Finkelstein, A. V. (2003) *Protein Sci.* **12,** 2057–2062.
42. Koga, N. & Takada, S. (2001) *J. Mol. Biol.* **313,** 171–180.
43. Lindberg, M. O., Tangrot, J., Otzen, D. E., Dolgikh, D. A., Finkelstein, A. V. & Oliveberg, M. (2001) *J. Mol. Biol.* **314,** 891–900.
44. Fersht, A. R. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1525–1529.
45. Shoemaker, B. A., Wang, J. & Wolynes, P. G. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 777–782.
46. Munoz, V. & Eaton, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11311–11316.
47. Alm, E. & Baker, D. (1999) *Curr. Opin. Struct. Biol.* **9,** 189–196.
48. Galzitskaya, O. V. & Finkelstein, A. V. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11299–11304.
49. Hubner, I. A., Shimada, J. & Shakhnovich, E. I. (2004) *J. Mol. Biol.* **366,** 745–761.
50. Krantz, B. A., Srivastava, A. K., Nauli, S., Baker, D., Sauer, R. T. & Sosnick, T. R. (2002) *Nat. Struct. Biol.* **9,** 458–463.
51. Wolynes, P. G. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6170–6175.