Scanning Malleable Transition State Ensembles: Comparing Theory and Experiment for Folding Protein U1A[†]

Tongye Shen,[‡] Christoph P. Hofmann,[§] Mikael Oliveberg,[⊥] and Peter G. Wolynes*[,‡]

Department of Chemistry & Biochemistry, University of California at San Diego and Center for Theoretical Biological Physics, La Jolla, California 92093-0371, Facultad de Ciencias, Universidad de Colima, Bernal Díaz del Castillo 340, Colima, Colima, Mexico, and Department of Biochemistry, Umea University, 901 87 Umea, Sweden

Received January 4, 2005; Revised Manuscript Received March 2, 2005

ABSTRACT: Using a variational free energy functional, we calculate the characteristics of the transition state ensembles (TSE) for the folding of protein U1A and investigate how they respond to thermal and mutational changes. The functional directly yields predicted chevron plots both for the wild-type protein and for various mutants. The detailed variations of the TSE and changes in chevron plots predicted by the theory agree reasonably well with the results of the experiments. We also show how to visualize the folding nuclei using 3D isodensity plots.

While a consensus is emerging on the main features of the mechanisms of natural protein folding, there remain some rather general questions that still need to be resolved. A question plaguing experimental studies is to what extent the folding transition state is malleable rather than fixed (*1*, *2*), that is, is it better to think of the transition state ensemble (TSE)[1] as a collection of structures that is "movable" by energetically perturbing along a reaction coordinate? Or should we ascribe to the TSE some rather configurationally well-defined specific states? Experimental studies of folding transition states require great care in execution and delicate finesse in interpretation (*3*, *4*). Meanwhile computer simulations of protein folding have already provided some vivid pictures and word descriptions of folding transitions (*5*, *6*). Yet because of worries about the accuracy of the force field and sampling problems, simulators often shy away from making explicit quantitative predictions of laboratory folding kinetics. The small free energy differences detectable by experiments through protein engineering kinetic studies are very hard to resolve computationally given the statistical sampling errors of current simulations. Direct routes to calculating free energy profiles on the other hand can resolve the small free energy changes that are measurable in the laboratory. While such calculations must make approximations, leading to, perhaps, systematic errors, their lack of statistical errors gives an advantage over simulations in many cases. While direct free energy evaluations do not have the burden of the sampling problem confronting simulations, they can still provide the microscopic description of the TSE in even greater detail than a typical kinetic experiment can access. One seldom stressed but important advantage of the free energy functional based theories over simulations is that the approach simultaneously treats unfolded ground state effects, which are often present in experiments owing to partial residual structure in the denaturant state, which may require carefully calibrated studies (*7*). Simulations have even greater problems fairly sampling the entropy dominated unfolded states than they do for the transition state ensembles.

To illustrate the power of free energy functionals to address transition state malleability, we examine the free energy profiles of protein U1A with a variational method developed by Portman, Takada, and Wolynes (*8*−*10*). The PTW variational method expresses the conformational free energy of a protein using a small set of constraining variables $\{C_i\}$ numbering not more than the number of residues. A reference trial Hamiltonian, $\beta\mathscr{H}_v = \beta\mathscr{H}_{ch} + 3/(2a^2)\sum C_i(\vec{r}_i - \vec{r}_i^*)^2$, defines these constraining variables that reflect the local Debye−Waller factors for main chain motions. Here $\vec{r}_i^*$ indicates the native position of residue i and $\mathscr{H}_{ch}$ is the backbone Hamiltonian. Using these variables, the free energy profile is easy to access and relatively smooth. The starting point of the PTW method is the (nonlocal) Hamiltonian of the protein that only contains the native contacts $\sum_{[ij]}$ and that uses the native structure as input information, that is, the target Hamiltonian is $\beta\mathscr{H}_t = \beta\mathscr{H}_{ch} + \beta\mathscr{H}_{int}$. The contact energy is modeled by a pairwise potential, $\mathscr{H}_{int} = \sum_{[ij]}u(|\vec{r}_i - \vec{r}_j|)$ with $u(r) = \sum_{k=s,i,l}\gamma_k \exp[\{-3/(2a^2)\}\alpha_k r^2]$. The parameters γ_k and α_k come from a tuning procedure (*9*). This starting model has a perfect funnel landscape (*11*−*13*). In such a simplified landscape, the interplay between stability and entropy alone determines the TSE. Of course, the real TSE will contain effects from nonnative interactions (*14*) and nonadditivity of contact formation energies (*15*). Since it is not convenient to calculate the partition function of $\mathscr{H}_t$, we use the Feynman−Gibbs−Peierls−Bogoliubov variational principle based on the reference Gaussian Hamiltonian, $\mathscr{H}_v$, to obtain the free energy landscape as a function of the constraint parameters $\{C\}$, that is, $F[\{C\}] \simeq -k_BT \ln Z_v +$

[†] Funding was provided by NIH Grant R01 GM044557.

* Corresponding author. E-mail: pwolynes@chem.ucsd.edu. Phone: (858) 822-4825. Fax: (858) 822-4560.

[‡] University of California at San Diego and Center for Theoretical Biological Physics.

[§] Universidad de Colima.

[⊥] Umea University.

[1] Abbreviations: TSE, transition state ensemble; PTW, Portman−Takada−Wolynes; MSD, mean-square deviation.

10.1021/bi0500170 CCC: $30.25
Published on Web 04/06/2005

$\langle \mathcal{H}_t - \mathcal{H}_v \rangle_v$, where Z_v is the partition function of the reference Hamiltonian.

The experimental manifestation of TSE malleability shows up in the so-called "chevron plots" for various mutations (*3*). For reversible two state folders, fast mixing and relaxation experiments yield directly the sum of the forward and backward rates, that is, $k_f + k_u$. The labels f and u refer to the dynamic processes of folding and unfolding, respectively. Commonly one plots the logarithm of the relaxation rate sum, $k_f + k_u$, against the denaturant concentration, c. At low denaturant concentrations, the folding rate, k_f, dominates the relaxation process, while k_u dominates at high denaturant concentrations. The resulting often stylized V-shaped curve is the so-called "chevron plot". If the TSE is not malleable, the chevron plot appears to be intersecting straight lines. But generally the plots do show curvature, which implies that either one or both of the transition state ensemble and denatured ensemble (or even the native state) are malleable (*16*, *17*). Curved chevron plots can also be the results of the folding intermediates. This is often observed in the case of larger proteins. Our small protein U1A clearly shows simple two-state behavior from exponential relaxation dynamics. There is no evidence of populated folding intermediates in this case. Our calculations do show fine structures of the free energy profile, which could be termed "high free energy intermediates".

The protein that we chose for our present study is U1A. U1A consists of a four-stranded antiparallel β-sheet with two α-helices packed on one side (*18*). It is one of the most popular RNA-interacting motifs, and it has been the subject of many experimental studies (*19*−*22*). For a theoretical analysis of transition state malleability, U1A offers a particular advantage by displaying characteristic nonlinearities/curvatures of the chevron plot that change both upon mutation and upon destabilization by chemical denaturants.

FREE ENERGY PROFILES AND FOLDING OF WILD-TYPE U1A AND ITS MUTATIONS

Being based on a native contact only Hamiltonian (Gō model), the PTW variational method uses the native contact information as input (*9*). While the overall picture of folding processes is not sensitive to the details of the definition of the contact map (such as cutoff distance, whether to use C_α atoms, C_β atoms, or any atoms of heavy-side chains, etc.), the fine structure of the free energy profiles depends to some extent on the precise definition of a native contact.

In Figure 1, we compare our results for the pathways obtained with two different definitions of the contact map: (i) C_β's being within a cutoff distance of 6.5 Å of each other and (ii) heavy atoms of side chains being within 4.0 Å, supplemented by some additional contacts found in the table of ref *22*. The definition (i) gives a sparse set of contacts with a folding transition temperature for wild-type U1A $T_f = 1.578$, while for definition (ii) there are more contacts leading to a higher $T_f = 2.033$. Note that, throughout the current work, we express energy and temperature in units of kcal/mol. Since the contact definition (ii) is consistent [while (i) is not] with the reported residue contact table for mutations designed in ref *22*, we focus on the results from definition (ii) below. The smaller the number of mean contacts per residue is, the more delicate and complicated

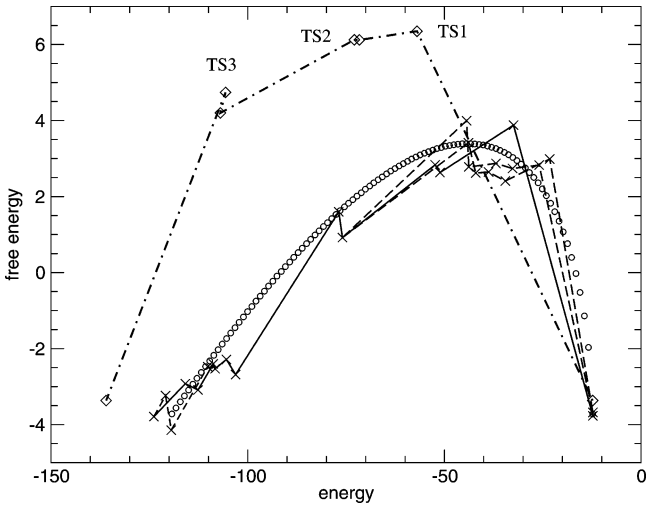

FIGURE 1: The wild-type folding pathways at T_f with backbone contact map (definition (i), curves with crosses, $T_f = 1.578$ kcal/(mol·R)) and side chain contact map (definition (ii), dotted-dashed curve with squares, $T_f = 2.033$ kcal/(mol·R)). The circle line, based on the capillary model, serves as a comparison with the results using backbone contacts.

the landscape is. The larger number of contacts generally increases the self-averaging and makes the contact network more robust against mutation and more cooperative, that is, it gives a somewhat higher transition barrier and a simpler transition pathway. Since the folding process is a battle between downhill lowering energy and decreasing entropy, we use for convenience the contact energy as a global reaction coordinate. It is approximately proportional to the average number of native contacts formed. When there is a fine structure along the folding profile, local reaction coordinates may be introduced and have already been studied with this method as highlightened by Portman et al. (*10*). As shown in Figure 1, the free energy profile obtained with the smaller number of contacts based on definition (i) agrees very well with the capillary calculation, which gives a profile $F(E) \approx ((-aE)^{2/3} - (-aE))$ and which also provides a parameter-free prediction of the peak position, $(E^{\ddagger} - E_D)/(E_N - E_D) = {}^8/_{27}$ (*23*). Here D and N label the denaturant and native state, respectively. The dense contacts, based on definition (ii), also shift the peak position $\beta^{\ddagger}$ somewhat to a more native position.

Occasionally the calculated folding path has "backtracking" when the path includes a part with a transition state that has fewer contacts than both the minima it connects (i.e., where the two unstable directions of the saddle point on the free energy landscape lead to.) In these cases, the protein has to unfold a bit before proceeding to fold to the native structure. This is an example of "topological" frustration. In the present case of U1A, the backtracking arises from two contact regions: (1) the turn connecting to the N-terminus of α1 (residues 20−22) with the turn connecting to the C-terminus of β2 (residues 45−48) and (2) the turn connecting to the C-terminus of β2 (residues 44−47) with the turn connecting to the N-terminus of β3 (residues 52−55). Similar results were also obtained for protein S6 with the same folding topology.

The folding routes for wild-type U1A and its mutants for several different temperatures are shown in Figures 2 and 3, respectively. On the basis of these profiles, we conclude that

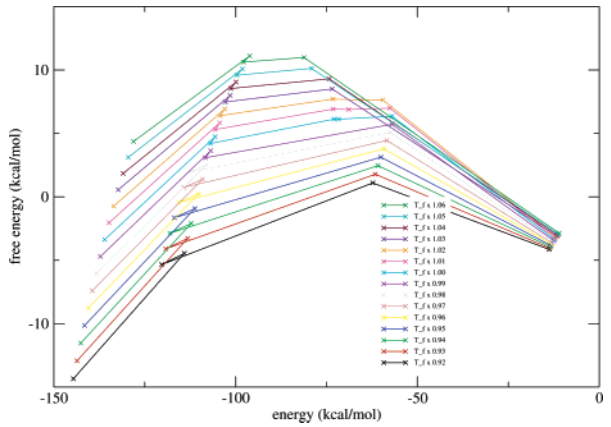

FIGURE 2: The wild-type folding pathway at different temperatures ($[0.92-1.06] \times T_f^{wild}$).

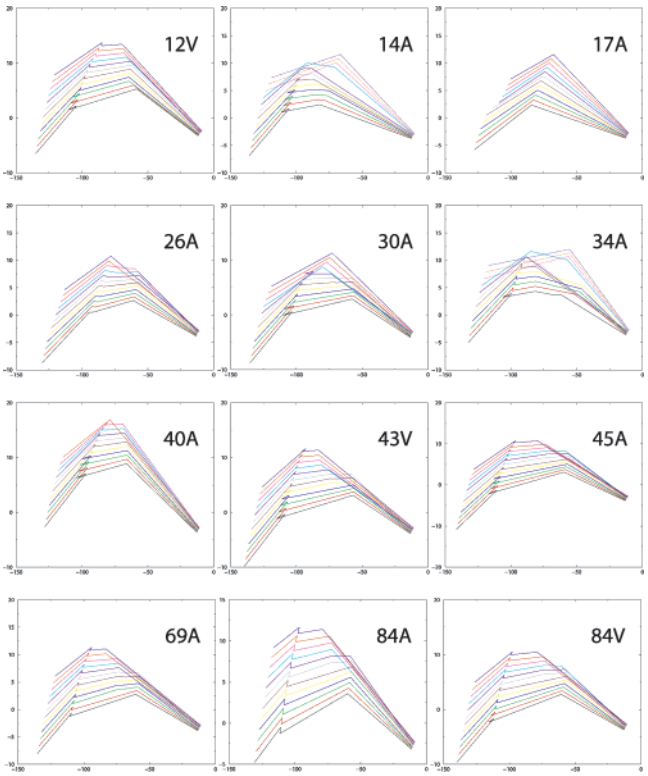

FIGURE 3: The folding pathways of mutants at different temperatures ($[0.94-1.05] \times T_f^{wild}$).

U1A and its mutants generally should show quite good two-state behavior without any significant population of intermediates.

Figure 4 shows how the position of the barrier, $\beta_E^{\ddagger} = (E^{\ddagger} - E_D)/(E_N - E_D)$, moves with changing the temperature. Here $E^{\ddagger}$ is the contact energy of the transition state with the highest $F^{\ddagger}$. It is clear from Figure 4 as well as from the routes shown in Figure 3, that individually most of the profiles exhibit signs of malleability. The locations of individual bottlenecks do not change much with denaturant, but relative barrier heights do change. In particular, the height of the dominant barrier changes easily with the change of the temperature. Typically, at low temperatures, crossing an early transition state is rate-limiting, but gradually the barrier from late transition states grows faster and becomes dominant with

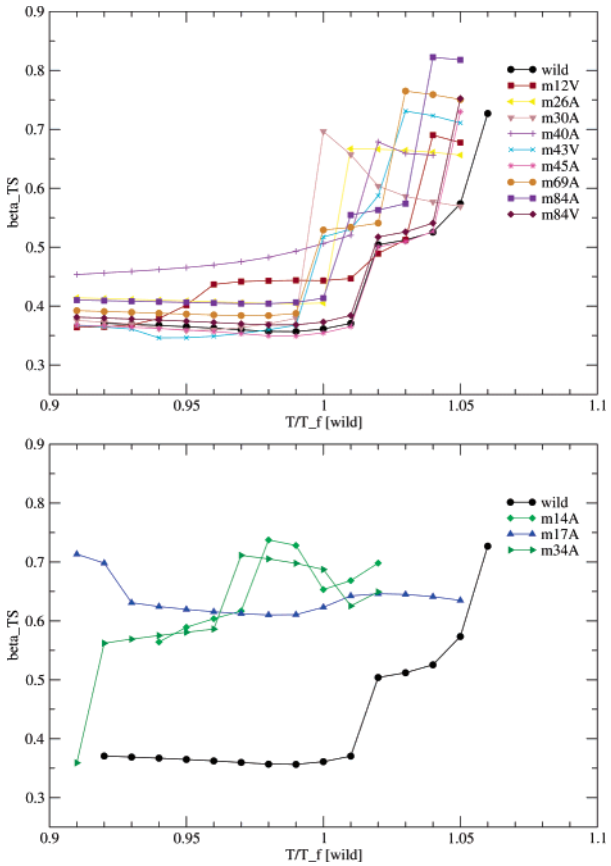

FIGURE 4: The movement of transition states with changing temperatures. Most of the mutants are quite movable. Some mutants (I14A, L17A, I34A) with a high β value move less than others.

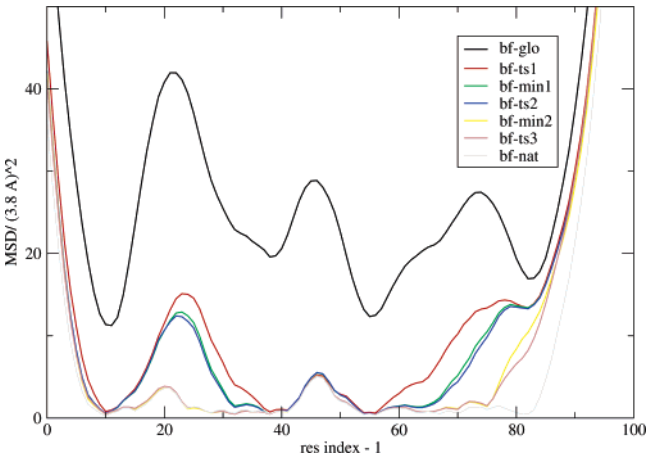

FIGURE 5: The mean-square-deviation plot for different states of wild-type U1A at T_f along the folding route. The unit is scaled by a factor of (3.8 Å)^2.

increasing temperatures. However, some mutants, such as L17A, show relatively more fixed transition states.

A more microscopic structural interpretation of the TSE can be made by examining the mean-square deviation (MSD) of each residue as predicted by the variational algorithm. Here the MSD of residue i is defined as how much residue i fluctuates around its mean position in structures belonging to the TSE, that is, the diagonal ith element of the covariance matrix $\mathbf{M}$. Here $M_{ij} = \langle(\vec{r}_i - \langle\vec{r}_i\rangle_v)(\vec{r}_j - \langle\vec{r}_j\rangle_v)\rangle_v/a^2$ and a = 3.8 Å. The MSD as a function of residue number is shown in Figure 5. Similar information is contained in the

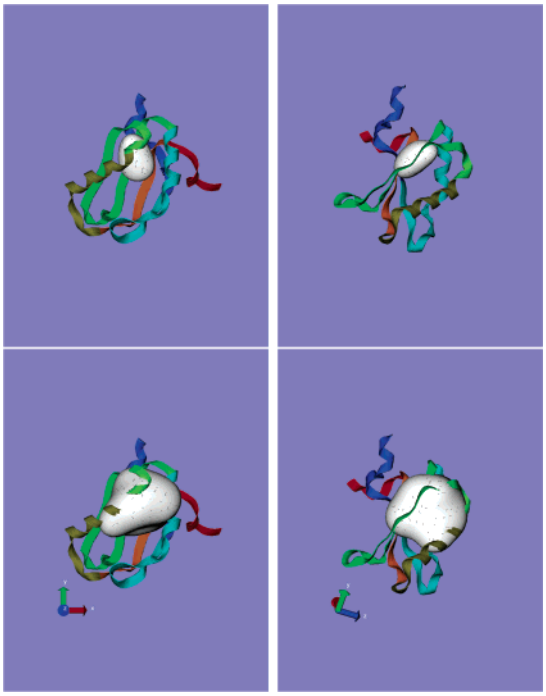

FIGURE 6: The isodensity surface ($\rho = 0.005$) at the early and late transition states for wild-type U1A at T_f. The early (top panel, two views) and late (bottom panel, two views) transition state are TS1 and TS3, which are labeled in Figure 1 and in the previous figure.

isodensity plot shown in Figure 6. The density field surrounding the protein is obtained from $\rho(\vec{r}) \approx \sum_i(\sqrt{M_{ii}^{-1}})^3 \exp(-3/(2a^2)M_{ii}^{-1}(\vec{r}_i - \langle\vec{r}_i\rangle_v)^2)$, where M_{ii}^{-1} is the localization parameter for residue i. They are obtained directly from the variational calculations for each state. A high number density (of residues) region will correspond to a more solid and more completely formed part of the protein. By displaying a surface of the same density, we can view the contour of the region corresponding to a classic folding nucleus. In Figure 6, we can see how the folding nucleus grows along the pathway, since different states along the folding path will give different density fields. Relatively speaking, the early transition state is less structured and has a small region of density higher than the contour surface of the specified fiducial density value; the later transition state is more formed and has a larger contour surface of the same density value. In this case, the delocalized nucleus centered around β-strand 3 and its connection with strand 2 and strand 1. The overall picture is consistent with the gestalt obtained from the experiments (*22*).

KINETICS AND CHEVRON PLOTS

Once we have free energy profiles, we can calculate the folding and unfolding rates and free energy−kinetics relations. Here we will approximate the folding rate for U1A by estimating the rate of crossing the TSE with the largest free energy barrier. As we have already noted, we do not find any deep intermediate states: the free energy profile is two-state-like with one broad barrier in the big picture. This one-barrier approximation makes our rate calculation straightforward. If the routes, however, were rougher and clear intermediate states were present, we would need to use other ways of estimating the relaxation rates such as the use of steady-state approximations. Steady-state approximations allow us to treat a multistep first-order reaction network as a single effective forward and backward rate process, but the steady-state approximation introduces its own problems. It does not respect the detailed balance principle, and one may obtain different rates k_f and k_u even for cases only having a slightly asymmetric pathway at the condition $G_N = G_D$. Note that we use the Gibbs (constant pressure) free energy G from here on since Gibbs free energies are experimentally measured whereas all the theoretical calculations are based on the Helmholtz (constant volume) free energy, F. We do not distinguish between these two in the current context since the volume change is small.

Also the steady-state rates are sensitive to the finer details of the pathway, making the comparison of results from different approximations to the pathways difficult. Our one-barrier approximation guarantees that $k_f = k_u$ for $G_N = G_D$, and its more robust behavior eases the comparison of results referring to different temperatures and point mutations. To estimate rates, we use a constant dynamical prefactor and thus ignore changes of rates due to the changing shapes of the basins and the barriers.

We compare our calculated rates with the experimental chevron plots, where the logarithm of the sum $k_f + k_u$ is plotted against the denaturant concentration c, that is, $P(c) = \ln[k_f(c) + k_u(c)]$. If we assume Arrhenius kinetics, $k(c) = A \exp[-\Delta G(c)/(k_B T)]$, with a constant prefactor A, we have

$$P(c) = \ln[\exp(-[G^{\ddagger}(c) - G_D(c)]/(k_B T)) + \exp(-[G^{\ddagger}(c) - G_N(c)]/(k_B T))] + \text{const}$$

In experimental studies, the chevron plots are used to derive primarily three parameters: (1) protein stability, (2) transition state position along the reaction coordinate, $\beta^{\ddagger} = m_f/m_{D-N}$, and (3) ϕ values for mutants.

1. Protein Stability. One thing to remember is that the concentration of the minimum point, c_x, of the chevron curve $P(c)$ is not precisely the concentration of unfolding, c_f (which we define by $G_N(c_f) = G_D(c_f)$). Experimentally c_f is often called the midpoint concentration, MP, which is obtained geometrically by the interception point of the linear extrapolations of the two wings of the chevron plot. We can estimate the relation between c_f and c_x by linearizing the effects of concentrations close to c_f with $c_x = c_f + \delta c$. Here δc satisfies the equation

$$(\partial/\partial c)[\exp -(b_D \delta c + \Delta G_o/(k_B T)) + \exp -(b_N \delta c + \Delta G_o/(k_B T))] = 0$$

where $\Delta G_o = G^{\ddagger}(c_f) - G_D(c_f)$ is the barrier height at c_f. This equation gives $\delta c = c_x - c_f = (b_N - b_D)^{-1} \ln(-b_N/b_D)$. We have defined $k_B T b_{D,N} = (\partial/\partial c)(G^{\ddagger} - G_{D,N})|_{c=c_f}$. We have generally $b_D > 0$ and $b_N < 0$ indicating higher folding barriers and lower unfolding barriers with increasing denaturant concentrations. We have zero shift, $\delta c = 0$, when $b_D + b_N = 0$. For earlier transition states, G_D is more likely coshifting with $G^{\ddagger}$ rather than G_N, and thus $|b_D| < |b_N|$, which means they have a tendency to exhibit negative (left) shifts,

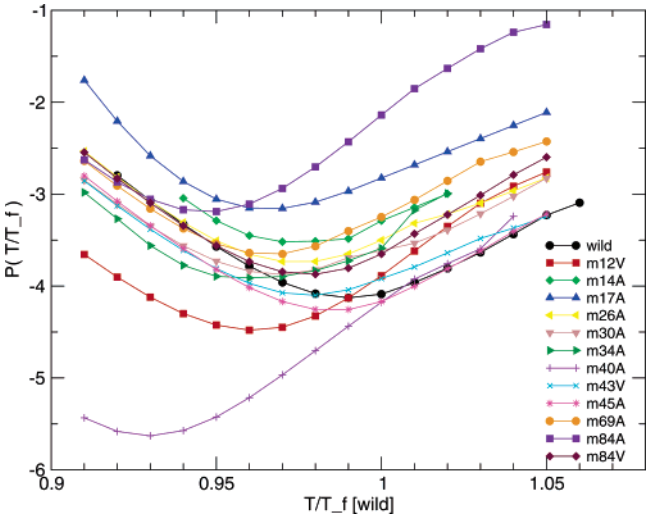

FIGURE 7: The thermal chevron plots of wild-type U1A and different mutants.

$c_x < c_f$, and vice versa positive (right) shifts, $c_x > c_f$, for late transition states.

The calculations presented do not directly deal with denaturant-induced unfolding; rather they probe the thermal denaturation curve:

$$P(T) = \ln[\exp(-[G^{\ddagger}(T) - G_D(T)]/(k_B T)) + \exp(-[G^{\ddagger}(T) - G_N(T)]/(k_B T))] + \text{const}$$

As in the case of the chemical denaturation, the midpoint temperature T_x (the temperature of the lowest point of the thermal chevron plot) and T_f (the folding temperature defined by $G_N(T_f) = G_D(T_f)$) will generally be different. Note that there is a temperature dependence in the denominator as well. After a similar calculation as for the chemical cases, the relation between T_x and T_f is

$$T_x = T_f\left[1 - \frac{1}{d_N - d_D}\ln\left(-\frac{d_N T_f + \Delta G_o}{d_D T_f + \Delta G_o}\right)\right]^{-1}$$

Here $\Delta G_o = G^{\ddagger}(T_f) - G_D(T_f)$ is the barrier height at T_f. Likewise, we define $k_B d_{D,N} = (\partial/\partial T)(G^{\ddagger} - G_{D,N})|_{T=T_f}$ with $d_D > 0$ and $d_N < 0$ as before.

In Figure 7, we show the thermal (calculated) chevron plots of wild-type U1A and its mutants. If a mutation shifts a chevron plot to the left of the wild-type, it means the protein will be less stable; if it shifts to the right of the chevron plot of the wild-type, it means the protein will be more stable. As we can see, the minima of all of the mutants that we chose (designed to delete contacts) shift to the left of the wild-type minimum indicating a decrease in protein stability.

In Figure 8, we also plot the chevron curve with translational movements of the individual curves shown above by centering the points at the folding temperatures of the individual mutations, T^{mut}, that is, $T \rightarrow T - T_f^{mut}$ and $P(T) \rightarrow P(T) - P(T_f^{mut})$. It can be seen that most curves have their minima T_x at slightly lower temperatures than their folding temperatures and three mutants, I14A, L17A, and I34A, have minima at slightly higher values. These observations are expected from our analysis and are consistent with the information from Figure 4 (part B) that late transition states have $T_x > T_f$. Practically, we found that for most

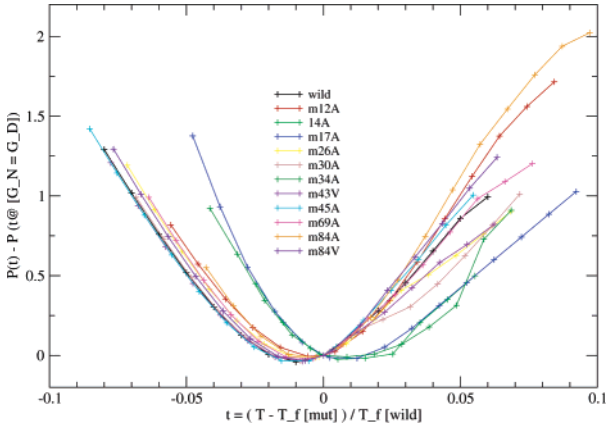

FIGURE 8: The shifted thermal chevron plots of wild-type U1A and different mutants by adjusting the origin of the plot with mutant T_f and its corresponding value (see text).

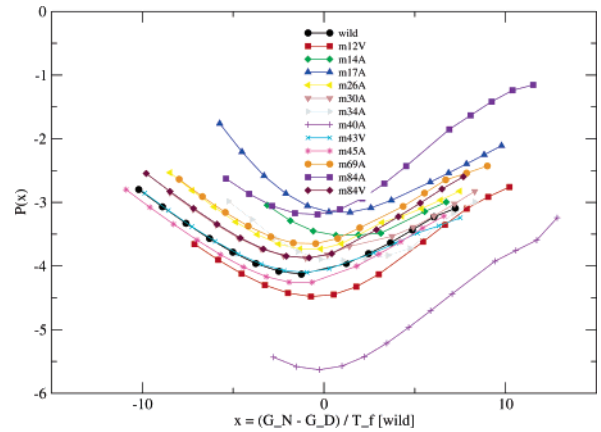

FIGURE 9: The thermal chevron plots of wild-type U1A and different mutants as functions of the corresponding ΔG_{DN} values.

mutants $|T_x - T_f|/T_f$ has a value of about 1%. We also plot the chevron values as a function of the stability ΔG_{DN} in Figure 9.

For a comparison between the calculations and experiments, we plot the experimental midpoint concentrations versus the folding temperatures of theoretical chevron curves in Figure 10. The correlation coefficient including all points is 0.73. All the correlations listed here and below are achieved without any tuning of the parameters.

2. Transition State Position $\beta^{\ddagger}$. Instead of the microscopic definition $\beta_E^{\ddagger} = (E^{\ddagger} - E_D)/(E_N - E_D)$ given in the section above and shown in Figure 4, the experimental values for $\beta^{\ddagger}$ are estimated from the ratio of the slopes of the chevron plot. In Figure 11, we plot the theoretical $\beta^{\ddagger}$ also obtained from the slopes of the calculated chevron plot near the transition region versus the corresponding results from experiments at the midpoint region. Without the clear outlier I40A, the correlation coefficient is 0.85.

3. ϕ *Values.* Though free energy perturbation for idealized chevrons allows a uniquely well-defined value for $\phi = \Delta\Delta G_{\ddagger D}/\Delta\Delta G_{ND} = \log(k_f^m/k_f^w)/\log(K^m/K^w)$, it is quite challenging to define ϕ values in a unique fashion for mutants of curved chevron plots ab initio. As detailed in ref *22*, rates from different concentrations for wild-type and mutants are used to calculate with the constraint of matching $\beta^{\ddagger}$. The alternate averaged ϕ^{mp} is defined by obtaining all the k_f at a fixed low concentration and k_u at a fixed high concentration

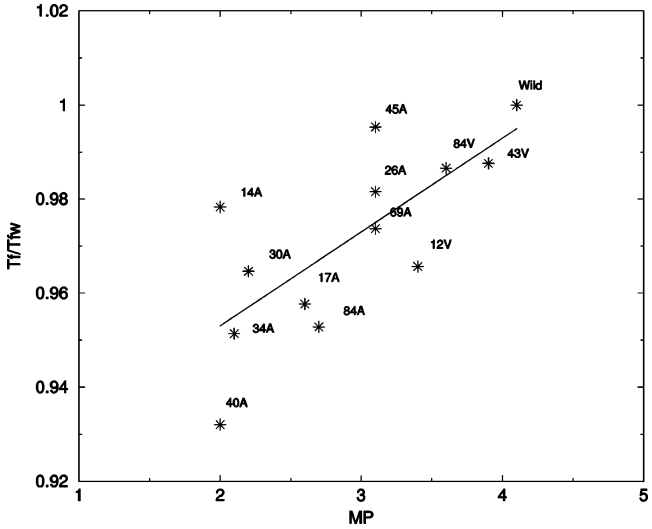

FIGURE 10: The comparison of experimental values of midpoint concentration vs computational values of T_f (normalized by T_f for wild-type) for various mutants.

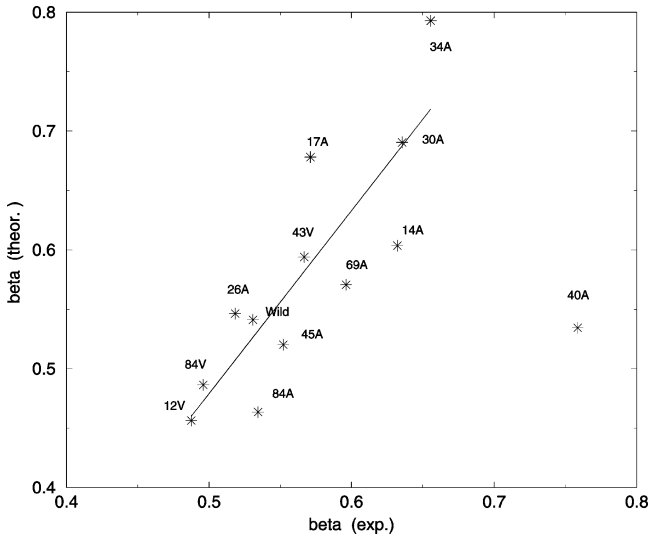

FIGURE 11: The comparison of experimental values of β at midpoint regions vs the corresponding values calculated from slopes of theoretical chevron curves.

of denaturants (*17*). We show in Figure 12 the comparison of ϕ values obtained by experimental curved chevron fits with theoretical calculations given by ϕ^{mp}. Here I43V is not shown in the figure since this mutation produces very small stability changes giving large experimental errors and thus leading to an apparent abnormal ϕ value larger than one. The corresponding theoretical ϕ for I43V is 0.662. We also did not show I14A and L17A in Figure 12. Both have negative ϕ's (−0.40 and −0.62, respectively) because the theory predicts faster folding rates k_f for mutants than the corresponding values for the wild-type, which are the results of the shift of barrier positions to relative late transition states. For L17A, this increase seems related to the fact that the mutations removed the backtracking. With the nine points shown in Figure 12, the correlation coefficient is 0.65. With all 12 points (including I43V, I14A, and L17A), the correlation coefficient is 0.66.

Finally in Figure 13, to make a direct comparison with experiments, we replotted the chevron data in a pairwise fashion between wild-type and mutants. We use the same

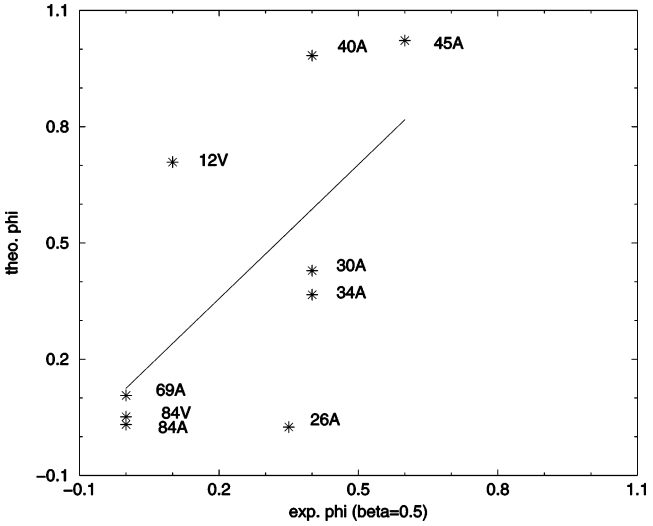

FIGURE 12: The comparison of experimental and theoretical ϕ values.

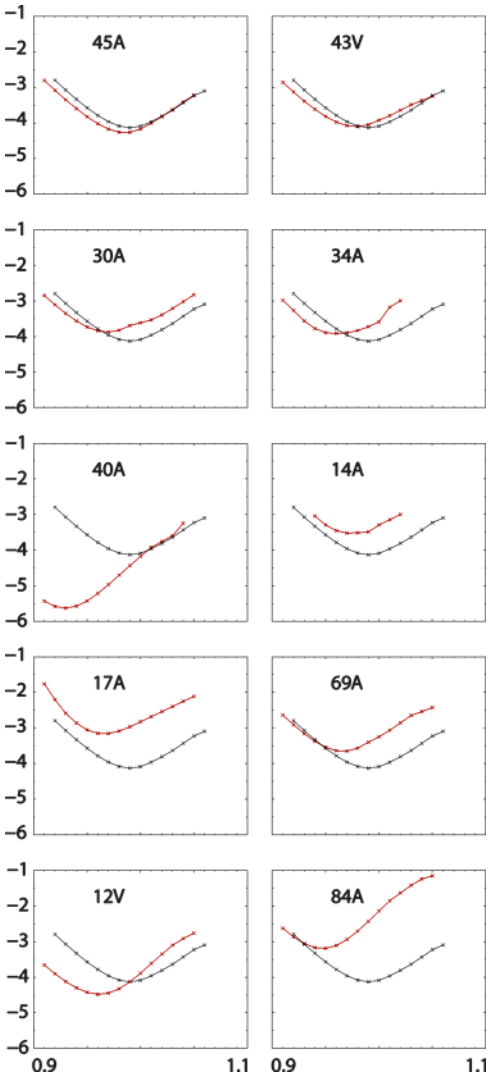

FIGURE 13: The thermal chevron plot comparisons. Mutation data are colored red; wild-type is colored black. They are listed in the same order as in Figure 3 of ref *22*. Note that the published experimental data are not shown here.

order of plots as in Figure 3 of the experimental work (*22*). For the majority of the mutants, experiment and theory agree reasonably well as to how mutations change the chevron plot

from that of the wild-type. The exceptions are I40A (which fails for both wings), I12V, and I14A (both fail on the left wing). For I12V, we have a slower folding wing than the experiment does, while for I14A, we have a faster folding wing than the experiment does. Still, the theoretical calculations of I40A show a high ϕ value, consistent with experiments. This mutant has the most serious loss of contacts found among all mutants, and thus finding deviations between the calculation (which is fundamentally a perturbation approach) and the experiment is understandable. I40A displays the largest distortions of the chevron slopes, which indicates that the barrier fine structure plays a critical role for the observed kinetics. The transition state barrier shift is larger for I40A than it is for other mutants. Also, unlike those for any other mutation, the calculations of $\beta^{\ddagger}$ from the two definitions give very different results for I40A. I12V is the most selective probe because it only involves one carbon and is thus more sensitive to local errors than the others.

CONCLUSIONS

We used the PTW variational method to investigate the malleability of the protein folding transition state ensemble for the well-studied system U1A. At a very coarse-grained level, the TSE emerges as "one" movable state with perturbations due to temperature, denaturant, and mutations. Zooming in, we can view it as arising from transitions among many fixed individual transition states. These states do not move significantly along the reaction coordinate β, but they do change their height (i.e., their relative free energy values) easily when the system is perturbed. Thus collectively they appear as one movable state. Because individual fine structures are sensitive to how the contact maps are defined, overall, we believe describing the system as having "one" movable TSE is a more robust formulation.

The overall gestalt is quite robust to the definition of the contact map. But we can learn in a perturbative sense from the results with a sparse contact map. The trend that we find from comparing the results from different definitions of the contact map is consistent with those that we have found in looking at other proteins (Hofmann et al., to be published): generally, having more contacts leads to a higher barrier and a more cooperative folding transition. Systems having more contacts are more likely to follow commonly observed two-state folding with a single broad barrier. On the other hand, when the systems have fewer contacts, they have a tendency to exhibit more parallel paths, some involving backtracking, and possibly the emergence of high free energy but visible intermediates. Proteins with sparse contact maps will have complicated folding routes.

Free energy functional methods allow one to calculate the chevron plot and compare the result directly with experiments. For U1A, this direct comparison shows that theory and experiment agree reasonably well and in a quantitative and not just qualitative fashion.

ACKNOWLEDGMENT

We thank Dr. J. J. Portman for providing us with the variational code and for helpful discussions. We also thank Dr. R. W. Hall and Dr. J. Gullingsrud for the help on the visualization of density fields with VMD. Computational support was provided in part by the Center for Theoretical Biological Physics.

REFERENCES

1. Yang, W. Y., and Gruebele, M. (2003) Folding at the speed limit, *Nature 423*, 193−196.
2. Sánchez, I. E., and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding, *J. Mol. Biol. 325*, 367−376.
3. Fersht, A. (1999) *Structure And Mechanism In Protein Science: A Guide To Enzyme Catalysis And Protein Folding*, W. H. Freeman, New York.
4. Nolting, B. (1999) *Protein folding kinetics*, Springer, Berlin.
5. Sato, S., Religa, T. L., Daggett, V., and Fersht, A. R. (2004) Protein folding funnels: A kinetic approach to the sequence-structure relationship, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6952−6956.
6. Wolynes, P. G. (2004) Latest folding game results: Protein A barely frustrates computationalists, *Proc. Natl. Acad. Sci. U.S.A. 101*, 6837−6838.
7. Sánchez, I. E., and Kiefhaber, T. (2003) Hammond behavior versus ground-state effects in protein folding: Evidence for narrow free energy barriers and residual structure in unfolded states, *J. Mol. Biol. 327*, 867−884.
8. Portman, J. J., Takada, S., and Wolynes, P. G. (1998) Variational theory for site resolved protein folding free energy surfaces, *Phys. Rev. Lett. 81*, 5237−5240.
9. Portman, J. J., Takada, S., and Wolynes, P. G. (2001) Microscopic theory of protein folding rates. I. Fine structure of the free energy profile and folding routes from a variational approach, *J. Chem. Phys. 114*, 5069−5081.
10. Portman, J. J., Takada, S., and Wolynes, P. G. (2001) Microscopic theory of protein folding rates. II. Local reaction coordinates and chain dynamics, *J. Chem. Phys. 114*, 5082−5096.
11. Gō, N. (1983) Theoretical studies of protein folding, *Annu. Rev. Biophys. Bioeng. 12*, 183−210.
12. Leopold, P. E., Montal, M., and Onuchic, J. N. (1992) Protein folding funnels: A kinetic approach to the sequence-structure relationship, *Proc. Natl. Acad. Sci. U.S.A. 89*, 8721−8725.
13. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes, *Science 267*, 1619−1620.
14. Viguera, A. R., Vega, C., and Serrano, L. (2002) Unspecific hydrophobic stabilization of folding transition states, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5349−5354.
15. Eastwood, M. P., and Wolynes, P. G. (2001) Role of explicitly cooperative interactions in protein folding funnels: A simulation study, *J. Chem. Phys 114*, 4702−4716.
16. Oliveberg, M. (2001) Characterisation of the transition states for protein folding: Towards a new level of mechanistic detail in protein engineering analysis, *Curr. Opin. Struct. Biol. 11*, 94−100.
17. Otzen, D. E., and Oliveberg, M. (2002) Conformational plasticity in folding of the split beta-alpha-beta protein s6: Evidence for burst-phase disruption of the native state, *J. Mol. Biol. 317*, 613−627.
18. Price, S., Evans, P. R., and Nagai, K. (1998) Crystal structure of the spliceosomal U2B″-U2A′ protein complex bound to a fragment of U2 small nuclear RNA, *Nature 394*, 645−650.
19. Allain, F. H.-T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation, *Nature 380*, 646−650.
20. Silow, M., and Oliveberg, M. (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
21. Silow, M., and Oliveberg, M. (1997) High-energy channeling in protein folding, *Biochemistry 36*, 7633−7637.
22. Ternström, T., Mayor, U., Akke, M., and Oliveberg, M. (1999) From snapshot to movie: phi analysis of protein folding transition states taken one step further, *Proc. Natl. Acad. Sci. U.S.A. 96*, 14854−14859.
23. Wolynes, P. G. (1997) Folding funnels and energy landscapes of larger proteins within the capillarity approximation, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6170−6175.

BI0500170