Acc. Chem. Res. **1998,** *31,* 765−772

Alternative Explanations for "Multistate" Kinetics in Protein Folding: Transient Aggregation and Changing Transition-State Ensembles†

MIKAEL OLIVEBERG‡

Biochemistry, Chemistry Centre, P.O. Box 124, S-221 00 Lund, Sweden

Received October 20, 1997

Introduction

Deviations from classical two-state kinetics in protein folding need not always be explained by the presence of rapidly formed intermediates. In some cases, such deviations are caused by short-lived aggregates whereas in other cases they arise from changes of the position of the transition state. These are two new facets of the mechanism of two-state folding. The first part of this account describes the effect of aggregates which form transiently in the first few milliseconds of the refolding reaction. The aggregates show many similarities with folding intermediates, but may be identified by their disappearance at low concentrations of protein where the two-state conversion of monomeric protein becomes predominant. In the second part, the focus is directed to two-state folding and movements of the transition state ensemble. The movements are used to derive information about the shape of the free-energy profile for folding. It emerges from a comparison of the kinetic behaviour of several small model proteins that the free-energy barrier for folding could be generally broad and level. An attractive feature of broad barriers is that, depending on minor variations in the fine-structure of the free-energy profile, they account for a wide range of seemingly unrelated folding data, including deviations from classical two-state kinetics determined by free-energy extrapolations.

1.1. Appearance and Role of Partly Structured States in Protein Folding

A characteristic feature of protein folding is the cooperativity which enable unfolded polypeptides to find their unique native conformations in one highly concerted step and without populated intermediates.[1] Although the cooperativity may thus be the single most important clue to the folding mechanism, it poses at the same time a great problem since the partly structured conformations which carry the details of this reaction are unstable and cannot readily be studied at equilibrium. Such all-or-none transitions are found for most small model proteins under equilibrium conditions, also those that are frequently reported to fold via populated intermediates.[2−5] But how does this make sense? Even if the intermediates cannot be found at equilibrium (an alternative and complementary approach to protein folding is to study proteins which deviate from the two-state rule and assume stable intermediate conformations at equilibrium under denaturing conditions[2,6]), they may accumulate transiently in time-resolved refolding experiments under stabilizing conditions prior to the formation of the native state (Figure 1). Before examining the role of these metastable intermediates, it is useful to divide them into two classes, depending on their origin: (1) kinetic intermediates that arise from nonnative restrictions in the denatured ensemble, for example, *cis*-proline isomers[6] and nonnative disulfide bonds;[7] (2) kinetic intermediates that arise in the folding process of small proteins without apparent restrictions in the denatured ensemble and which may be an integral part of the folding mechanism. It is the role of the latter class of kinetic intermediates which has been the subject of intense analysis and long-lasting controversy.[4,8−12].

An early view was that the collapse into a series of "on pathway" intermediates is needed to collect unfolded polypetides and guide them through the overwhelming space of possible conformations.[6,13] More recent ideas include the whole complexity of the conformational space and suggest that the polypeptide can converge to the native structure from a multitude of directions.[14−18] In this perspective, metastable intermediates might just represent circumstantial minima in the energy landscape and terms as "on pathway" or "off pathway" are avoided to emphasize the width of the delocalized ensemble of conformations which progress toward the single native state.

A major reason for questioning the role of kinetic intermediates is that a growing number of proteins have been observed to fold efficiently without them (ref 19 and references cited therein). Furthermore, these two-state proteins usually fold faster than proteins which accumulate intermediates. This has led to the view that, given a fixed transition state, populated intermediates are better avoided since they would only contribute to slow folding and lower the stability of the native state.[3,11] Instead, and in analogy with the principle for enzyme catalysis, an optimal folding rate is suggested to result if the nucleation of structure stabilizes selectively the transition state.[11]

Mikael Oliveberg was born 1963 in Göteborg, Sweden. He has a background in physics and biology and received his Ph.D. in 1992 on electron-transfer reactions and proton pumping in cytochrome *c* oxidase. His postdoctoral training was on protein folding at MRC in Cambridge, U.K. Since 1996 he runs his own folding group at the Department of Biochemistry at Lunds University, Sweden. Mikael Oliveberg was appointed docent in 1997.

† Abbreviations: D, denatured ensemble; I, partly structured ensemble "folding intermediate"; ‡, transition state ensemble; N, native ensemble.

‡ Telephone: +46 46 222 01 07. Fax +46 46 222 45 34. E-mail: mikael.oliveberg@biokem.lu.se.

10.1021/ar970089m CCC: $15.00
Published on Web 08/29/1998

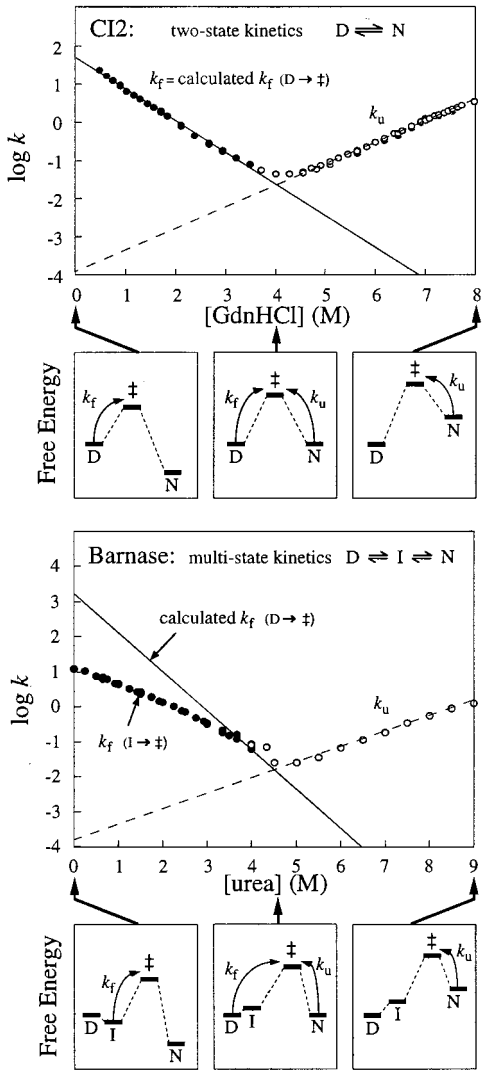

FIGURE 1. Chevron plots and free-energy diagrams for CI2[23] and barnase,[3] which constitute the classical examples of two-state and multistate kinetics. The lines are two-state folding calculated from eq 1. With CI2, the observed refolding rate constant (k_f) matches precisely that calculated from eq 1, indicating that refolding takes place directly from D. With barnase k_f is lower than the calculated refolding rate constant at low [urea], which is taken as evidence for accumulation of a dead-time intermediate.

1.2. Practical Aspects: How To Generate Kinetic Intermediates

Time-resolved refolding experiments are usually done by rapid dilution of chemically denatured protein with water in a stopped-flow apparatus[20] or in a quench-flow setup for the production of NMR samples.[21] The time required to mix the buffers (milliseconds), however, is too long to resolve the formation of the transient intermediates, which is rapid and takes place before the sample reaches the observation chamber. The observed reaction is thus the conversion of the intermediate into the native protein. Since the lifetime of the intermediates is relatively short (seconds), they are difficult to characterize and very little is known about their structure or dynamic properties. It is even uncertain whether they represent true intermediates, which are separated from the denatured ensemble by an activation barrier, or just compact cross-sections of a denatured continuum.[4,15,22] The crux of studying early folding events is consequently to resolve the rapid collapse of the polypeptide under stabilizing conditions, and significant progress toward the understanding of these fast processes is summarized in the current issue of *Accounts of Chemical Research*.

1.3. Identifying Kinetic Intermediates by Retardation of the Folding Rate

The most obvious way to identify kinetic intermediates is by spectroscopy, i.e., to see if the spectrum of the dead-time species differs from that of the denatured protein.[2,6] The approach discussed here is altogether different and based on the construction of free-energy profiles from kinetic data and protein stability.[3,5] According to the energy diagrams in Figure 1, folding directly from the denatured state is faster than from a downhill intermediate, because in the latter case the activation barrier is higher. This makes it possible to identify intermediates from the refolding rate. A minimal analysis is to see how the refolding kinetics changes with protein stability in a, so-called, chevron plot. In a two-state reaction,[23] the refolding and unfolding sides of the chevron plot are linear since the activation free energy depends linearly on [GdnHCl] (Figure 1). In the event that an intermediate becomes populated at low [GdnHCl], this retards refolding and causes a downward curvature of the chevron plot[3,58] (Figure 1). A more strict analysis is to extrapolate the unfolding kinetics (log k_u vs [GdnHCl]) into the refolding region and use this to calculate the refolding rate constant (k_f) according to the two-state relation[3]

$$K_{D-N} = [D]/[N] = k_u/k_f \Rightarrow \log k_f = \log k_u - \log K_{D-N} \quad (1)$$

where log K_{D-N} is derived from equilibrium denaturation experiments.[24] The calculated value of the refolding rate constant (log k_f^{calc}) is then compared with the observed value (log k_f^{obs}), and if there is a difference, this is taken as a deviation from two-state folding. It should be noted that the procedure assumes that the transition state is the same at all concentrations of denaturant. According to arguments developed in the second part of this account, this may not always be true. Based on a three-state model and assumptions of a fixed transition state, changes in log k_f^{calc} − log k_f^{obs} upon point mutations have been elegantly used to map out the structure and properties of the kinetic intermediate of barnase.[25] At present the barnase studies provide, by far, the most detailed characterization of a metastable intermediate, demonstrating the potential of this approach. However, care must be exercised when assigning the pathway based on deviations from two-state kinetics, as is demonstrated by the U1A protein.

1.4. Intermediates by Transient Aggregation

With the 102 residue U1A protein the deviation from two-state kinetics has proved to have a very different origin,

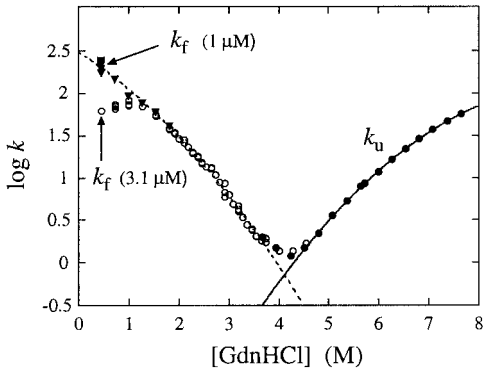

FIGURE 2. Chevron plot for U1A.[12] The fitted curves are second-order polynomials which obey eq 1 and indicate two-state folding D ↔ N. At 3.1 μM protein, the refolding of U1A is slower than expected from eq 1 (○), whereas at 1 μM folding is two-state (▼). The the retardation of k_f at 3.1 μM is caused by transient aggregation of D under refolding conditions, where the solubility of hydrophobic residues is poor.[12]

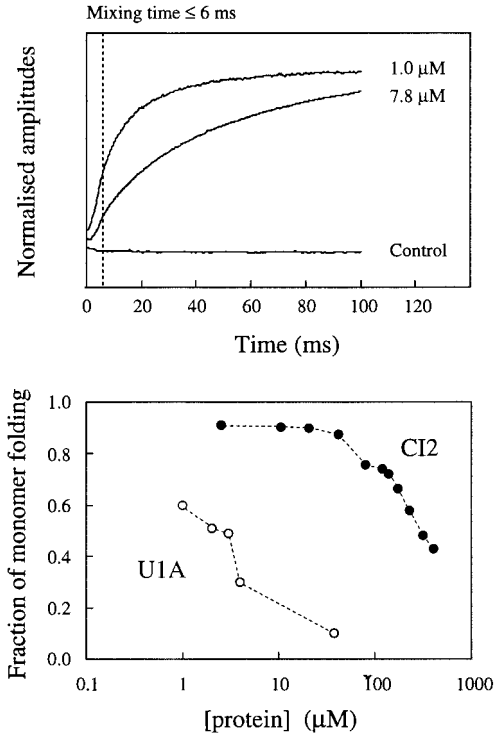

FIGURE 3. (Top panel) Time courses for refolding of U1A at different protein concentrations, showing biphasic kinetics at 1 μM and a single phase at 7.8 μM.[12] (Bottom panel) Fraction of monomer folding of U1A and CI2 at different protein concentrations, derived from the ratio of the amplitudes of the fast and slow refolding phase.[12]

namely, aggregation[12] (Figure 2). At moderate to high concentrations of protein (>5 μM), the refolding reaction is described solely by a slow phase (~40 s^{-1}), giving rise to a seemingly normal multistate behavior. At lower protein concentrations, however, higher-order events are revealed by the appearance of a faster phase near physiological conditions (Figure 3). Interestingly, the rate constant for this fast phase matches precisely that expected for two-state folding according to eq 1. From here, it is a short step to conclude that the slow phase is refolding from aggregates, and the fast phase, two-state

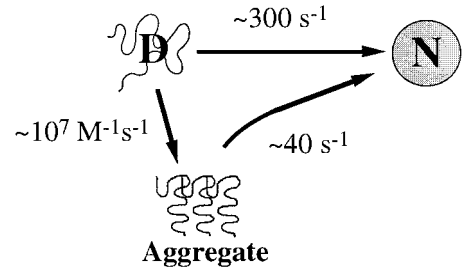

FIGURE 4. In vitro folding of U1A and CI2 in physiological buffer. At low concentrations of protein, refolding takes place directly from the denatured state, whereas at high concentrations of protein refolding takes place from short-lived aggregates. The rate constants are for U1A.

refolding of the monomeric protein (Figure 4). Notably, the rate constants for the two phases display only a limited dependence on protein concentration, and the aggregation behavior is revealed mainly by the amplitudes.[12]

Transient aggregates are easily mistaken for metastable intermediates since they are not easily revealed by changes in protein concentration under conditions where they are fully populated. With U1A the two-state folding is apparent only below 1 μM of protein. This means that U1A would be taken as a fairly "normal" three-state protein with an intermediate above ~5 μM (>0.06 mg/mL) (Figure 3). Notably, most stopped-flow CD and quench-flow NMR experiments use considerably higher protein concentrations and would never resolve the two-state characteristics of U1A. Transient aggregation under very dilute conditions is found also in disulfide-trapping experiments with α-lactabumin,[26,56] and more recently with phosphoglycerate kinase at a protein concentration as low as 0.05 μM (<10^{-3} mg/mL).[27] In other cases, the aggregation becomes a problem first at higher protein concentrations. For example with the two-state protein chymotrypsin inhibitor 2 (CI2), the aggregation is hardly noticeable at 5 μM but leads to multistate kinetics with an apparent kinetic intermediate above 200 μM (>2 mg/mL)[28] (Figure 3). It appears then that by simply varying the protein concentration, an apparent three-state folder can be turned into a two-state protein and, vice versa, a two-state protein into a three-state folder. A growing number of proteins seem to display this behavior.

1.5. Kinetic Partitioning into Fast and Slow Phases, Indicating That Aggregation Occurs in Competition with Folding

The biphasic refolding of U1A and CI2 suggests that the aggregates are not in dynamic preequilibrium with the denatured monomers, since this would give only one refolding phase in which the aggregates and monomers disappear concurrently.[29,30] A reasonable assumption is then that aggregation occurs in competition with folding, which also allows an estimate of the association constant (Figure 4). For U1A the rate constant for dimerization would be 4×10^7 M^{-1} s^{-1}, which is just below the diffusion limit.[31] At first this may appear unrealistically fast, but similar association rates are observed to control

the folding of the dimeric arc repressor.[32] For CI2 the association rate appears considerably slower at 3×10^5 M^{-1} s^{-1}, which may be explained by a higher net charge or differences in the conformation of the denatured states.[33] An immediate prediction is still that slowly refolding proteins or mutants would tend to aggregate more.

1.6. Determination of Which Species Gives Rise to the Transient Aggregates

Judging by the results from the two-state proteins U1A and CI2, it is the unfolded coil which forms the transient aggregates, whereas in most other studies aggregation is assigned to intermediates or molten globule conformations.[34,35] However, aggregation is often monitored over longer time scales and, sometimes, at very high protein concentrations where short-lived and reversibly formed multimers are easy to miss. It is therefore an open question if early events in macroscopic aggregation involve smaller complexes of more unfolded species as is observed with U1A and CI2, in particular at low concentrations of denaturant.

1.7. A Few Ways To Identify Short-Lived Aggregates

Some important properties of the transient aggregates are that (i) they are short-lived (ms) and, as with real folding intermediates, their conversion into the native protein is more or less a first-order process and (ii) the folding of monomeric protein may be resolved only at very low concentrations of protein (<2 μM) which are difficult to handle experimentally. Moreover, the folding of the monomeric protein could be very fast and lost in the mixing time.[36] Simple tests for transient aggregation are plots of refolding amplitudes against protein concentration and stopped-flow "double jumps". In the absence of aggregates, the refolding amplitudes should display a linear dependence on protein concentration and intersect with the origin. Likewise, refolding−delay−unfolding experiments should yield the expected unfolding kinetics and precise agreement with standard refolding experiments. It has also been revealing to do quench-flow labeling[21] at low protein concentrations (<1 μM[57]). The first and possibly most illustrative test, however, is to drip from a pipet a concentrated solution of denatured protein into a glass of water. If the solution becomes transiently opaque, aggregation is a problem for certain.

2.1. Barriers for Folding. The Possibility of Characterizing Two-State Reactions Where All Partly Structured States Are Unstable and Never Seem to Accumulate

The most direct way to probe free-energy barriers is by rate constants: the higher the barrier, the slower the protein will fold. Since we do not yet know what is causing the barrier, however, we are unable to quantify its height ($\Delta G^{\ddagger}$).[37−40] An ad hoc solution has been to use

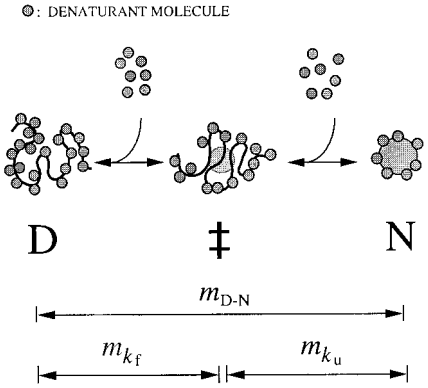

FIGURE 5. The meaning of *m*-values. From mass action, $m_{D-N} = \Delta Q_{D-N}/[GdnHCl]$, $m_{k_f} = \Delta Q_{D-\ddagger}/[GdnHCl]$, and $m_{k_u} = \Delta Q_{N-\ddagger}/[GdnHCl]$, where ΔQ is the number of GdnHCl molecules taken up or released in the respective transitions. The *m*-values are related to changes in the number of solvent accessible GdnHCl binding sites and used to determine the position of ‡ relative to D and N on a simple reaction coordinate based on solvent exposure[1] (eq 3).

transition-state theory, although this is derived for elementary reactions and overestimates $\Delta G^{\ddagger}$ for protein folding, where the reconfiguration time is several magnitudes slower than for small molecules.[41] Even so, it possible to relate changes in rate constants to changes in $\Delta G^{\ddagger}$ by assuming that the downhill rate (prefactor) is independent of the barrier height, denaturant, and amino acid composition.[28] This approach has been successfully used to map out the interactions in the transition state by systematic substitution of amino acids by protein engineering.[42−45] It appears from these studies that there are considerable differences in the transition-state ensembles, ranging from expanded structures with a low structural content for CI2[43] and Che Y[44] to a fairly compact and nativelike transition state for barnase.[42]

A more easily derived parameter is the position of the transition state on the D to N reaction coordinate (β). Typically this reaction coordinate is based on solvent exposure of the polypeptide which is derived from the so-called *m*-values [1,24,28,46] (Figure 5).

$$m_{D-N} = \frac{\partial \log K_{D-N}}{\partial[GdnHCl]}, \quad m_{k_f} = -\frac{\partial \log k_f}{\partial[GdnHCl]}, \quad m_{k_u} = \frac{\log k_u}{\partial[GdnHCl]} \qquad (2)$$

The proportionality between *m*-values and changes in solvent exposure of the polypeptide[1] makes it possible to estimate the position of the transition state according to

$$\beta^{\ddagger} = m_{k_f}/m_{D-N} = 1 - m_{k_u}/m_{D-N} \qquad (3)$$

where $\beta^{\ddagger}$ is a value between zero and one,[1] m_{D-N} is derived from equilibrium titration,[24] and m_{k_f} and m_{k_u} are from chevron plots.[28] Interestingly, $\beta^{\ddagger}$ reveal significant variations in the compactness of the transition-state ensemble for different proteins and, in several cases, also upon mutation or addition of denaturant.[36,47,48] An explanation for this variation arises from the symmetrically curved chevron plot of U1A.

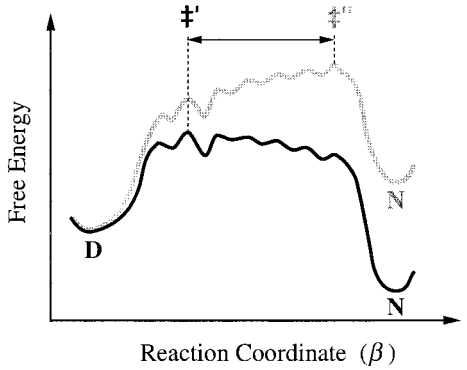

FIGURE 6. Movements of the transition-state ensemble (‡) on top of a broad activation barrier. The black graph illustrates the free-energy profile for folding under stabilizing conditions, and the gray graph illustrates the profile under denaturing conditions.

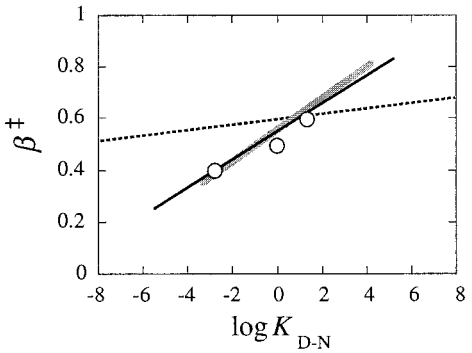

FIGURE 7. Position of the protein folding transition state ($\beta^{\ddagger}$) plotted against protein stability (log K_{D-N}). With U1A, the transition state moves from $\beta^{\ddagger} = 0.26$ in pure water to $\beta^{\ddagger} = 0.84$ in 8 M GdnHCl (black line). Consistent results are obtained from a 27-mer lattice model[39] (○). With wild-type CI2, the corresponding changes of the transition state are much smaller and cover only about 10% of the reaction coordinate (thin dotted line). Upon certain mutations, however, CI2 begins to display transition-state movements very similar to those of U1A (gray line).

2.2. Curved Chevron Plots, Suggesting Movements of the Transition-State Ensemble and Broad Activation Barriers

The curved chevron plot of U1A has been ascribed to a broad activation barrier for folding which causes the transition state to move toward the native state upon destabilization[19] (Figure 6). The evidence is mainly that $\beta^{\ddagger}$ increases upon addition of denaturant, while a constant m_{D-N} and the kinetic amplitudes indicate that D and N remain the same. The movement of the U1A transition state is comparatively large and covers more than 60% of the reaction coordinate[19] (Figure 7). It is important to realize that transition-state movements follow naturally from mass action and the way the reaction coordinate is constructed; i.e., a biased destabilization of the reaction profile upon addition of GdnHCl causes the barrier to tilt.

Broad barriers are found also by theory,[15,17,49] and Plotkin et al.[39] use a 27-mere lattice model to predict transition-state movements which are more or less identical with those of U1A (Figure 7). The first experimental reports on transition-state movements came from Matouschek and Fersht,[50] who observed effects on m_{k_u} upon mutation and discussed these in the perspective of the Hammond postulate in physical chemistry.[51] However, these movements are much smaller than those observed for U1A and give rise to relatively weak curvatures in the kinetics.[52] In terms of the shape of the free-energy profile, this would indicate a rather pointed barrier, although it must be realized that a fixed transition state shadows the rest of the activation barrier the shape of which is unknown. More pronounced movements are found by Sauer and co-workers for the late transition state of the dimeric arc-repressor, where $\beta^{\ddagger}$ varies between 0.69 and 0.92, depending on the mutation.[53] Like here, the authors suggest that the transition-state movements could be associated with a "high and bumpy free-energy plateau".[53]

Note that with respect to the rate-limiting step and the nature of the apparent transition state the broad barrier description seems too simplistic. A transition state needs to jut up some kT from the neighboring free-energy profile in order not to be blurred by the thermal motion. Without such distinct features, as will be the case with a smooth and level barrier, the reaction will tend to become diffusive[16] and the interpretation of $\beta^{\ddagger}$ is more complex. Nevertheless, the broad barrier description captures well the experimental behavior of protein folding, and it will be interesting to see how this issue will finally be resolved.

An alternative explanation for curves and changes in the unfolding kinetics is accumulation of unfolding intermediates,[59] i.e., the decreasing m_{k_u}-values are caused by partial disruption of the native structure in the dead time of the unfolding experiment. In the case of U1A, however, we have not yet found any good evidence for this scenario and choose therefore to focus on the effects of broad barriers which have been largely neglected.

2.3. Determining Whether or Not Broad Activation Barriers Are a General Feature of Cooperative Folding

In a recent comparison of several small proteins, it was found that broad activation barriers may not be particular to proteins displaying curved kinetics but could be a general feature in two-state folding[28] (Figure 8), this despite marked differences in the folding kinetics. A key observation is that certain point mutations of CI2 set free large movements of the transition-state ensemble, resulting in U1A like curvatures in the unfolding kinetics.[28] Another critical observation is that destabilization in some cases leads to sudden changes between two widely separated transition states, indicated by kinks in the chevron plots. The behavior is observed with arc-repressor, where, in addition to kinks, deletion of a salt bridge reveals two new transition states at other positions of the reaction coordinate[48] (Figure 8), and with a λ repressor fragment, where $\beta^{\ddagger}$ changes from 0.39 to 0.84 upon substitution of two alanines for glycines.[36] Taken together, these results show that seemingly large alterations of the free-energy profile are induced by relatively small perturbations of the protein. A simple explanation for this behavior is a broad activation barrier where localized

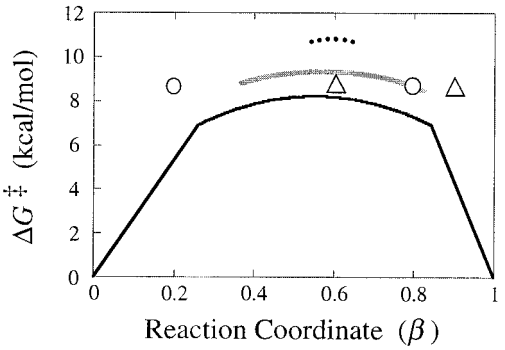

FIGURE 8. Comparison of the shapes of activation barriers at the transition midpoint. Barrier heights are calculated from transition-state theory with a downhill rate constant of 10^6 s^{-1},[41] and top surfaces are estimated by linear free-energy extrapolations.[28] The black profile illustrates the barrier of U1A. Further evidence for broad barriers are found with arc-repressor,[48] where two widely spaced transition states at $\beta^{\ddagger} \approx 0.6$ and $\beta^{\ddagger} \approx 0.9$ (△) change to $\beta^{\ddagger} \approx 0.2$ and $\beta^{\ddagger} \approx 0.8$ (○) upon mutation of a buried salt bridge. The barrier of wild-type CI2 appears relatively pointed (dashed black profile) but can be turned into a broad barrier by mutations (gray).[28]

transition states are minor features jutting up from the top.[28] In the case of CI2, the mutations with enhanced transition-state movements level out the isolated feature which normally constitutes the transition state and the reaction profiles become smooth as for U1A (Figure 9A,B). This leveling could either occur by lowering the feature or, perhaps more likely, by elevating its surroundings. Wild-type arc-repressor would then display two auxiliary features on top of the broad barrier and a free-energy profile which is overall more rugged. Under stabilizing conditions, the early feature is the transition state, whereas at high concentrations of denaturant the later feature becomes the transition state with a consequent change in the kinetic *m*-values [48] (compare Figure 9C,D). The mutation simply alters the fine structure of the free-energy profile so that other parts of the rugged barrier are exposed. Analogously the kink in the unfolding kinetics of cyt *c* [54] may indicate a change of transition state at 4 M GdnHCl (Figure 9C).

2.4. Changes of the Protein Folding Transition State, Providing Yet Another Two-State Explanation for Apparent Multistate Kinetics

An interesting situation occurs when the change of transition state occurs before or within the unfolding transition. By standard analysis this could still yield approximately linear free-energy relations since the *m*-value shifts may appear near the bottom of the chevron plot where they are hard to resolve.[48] However, extrapolation of unfolding kinetics from high [denaturant] would overestimate the unfolding rate constant in pure water since the transition state has changed (Figure 9D). This produces an apparent deviation from two-state kinetics which may be mistaken for a folding intermediate. Looking at stability and kinetics alone, it is thus easy to confuse changes of the transition state and the existence of kinetic intermediates, and the intermediates cannot be trusted

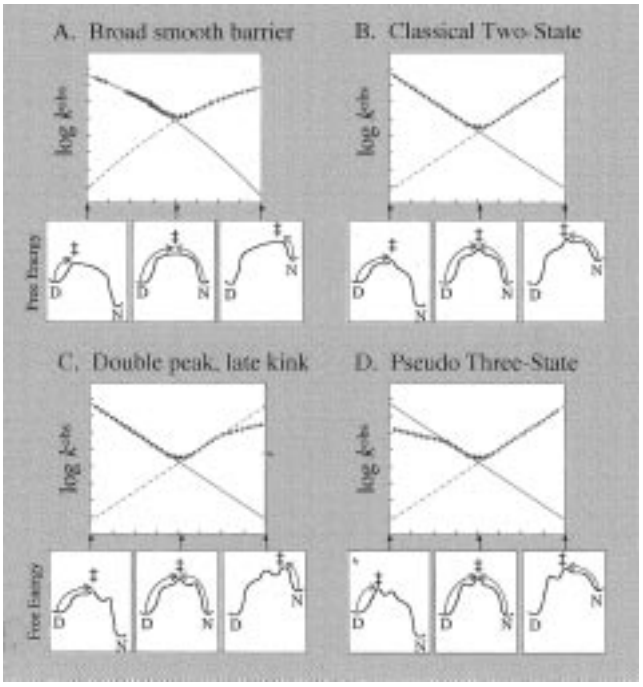

FIGURE 9. How broad activation barriers account for a wide range of kinetic data without the need to invoke populated folding intermediates. The color coded arrows show the activation process behind the observed kinetics, and the lines relate to eq 1. (A) A broad and smooth barrier causes large movements of ‡ which give rise to pronounced curvatures in the chevron plot. The behavior is observed experimentally with U1A[19] (Figure 2). (B) If a secondary peak juts up from the top of the barrier and constitutes the highest point at all [denaturant], the transition state will be fixed. This is manifested in constant *m*-values for the kinetics and a V-shaped chevron plot (compare CI2 in Figure 1). (C) If two pronounced peaks jut up from the barrier, a biased destabilization of the reaction profile may cause a sudden change between two well-separated transition states. The change is seen as a downward kink in the unfolding kinetics caused by a sudden decrease in m_{k_u} as the activation process for unfolding changes from an expanded (red) to a more compact (green) transition-state ensemble. (D) When the secondary peaks are located early in the activation process, the change of transition state (green to blue) may take place under stabilizing conditions. Equation 1 would here overestimate the refolding rate at [denaturant] = 0 (solid line) since the unfolding kinetics refers to a more compact transition state. If this deviation is assigned to a folding intermediate as in Figure 1, the apparent stability of the intermediate is the difference in free-energy between the two transition states.

unless verified with other techniques, for example, quench-flow NMR at low protein concentrations.

2.5. Determining Whether Folding Events Are Always the Same or If They Change upon Addition of Denaturant

It is easy to imagine that folding occurs in more than one way as the polypeptide can attain such a tremendous number of different conformations.[18] Denaturants which interact differently with the various species on the folding landscape are likely to perturb the free-energy profile, not only along the reaction coordinate but also "perpendicular" to the reaction coordinate.[55] This means that the

expanded transition-state ensemble identified at low [GdnHCl] may not be part of the folding reaction at [GdnHCl] = 8 M, or *vice versa*. Nor need an intermediate, albeit "on pathway" at low concentration of denaturant, remain "on pathway" as an unstable intermediate under two-state conditions at high [GdnHCl]. The problem is illustrated by the following hypothetical example: In pure water, a small nativelike helix forms independently and accumulates early in the folding reaction as a kinetic intermediate. The helix is then carried along like a satellite until it is integrated with the rest of the structure in the transition state. Upon addition of denaturant, the helix is destabilized relative to the coil, and since local interactions are no longer able to maintain its structure, it cannot form without support from the rest of the polypeptide. This happens first in the transition state. Accordingly, the folding sequence in pure water is the formation of local interactions followed by docking of elements of secondary structure (framework folding[2]), whereas in the presence of denaturant the formation of local interactions is driven by long-range interactions in the transition state (nucleation condensation[11]). Importantly, the denaturant has changed the folding pathway. The example illustrates how studies of the isolated helix are misleading since its accumulation is circumstantial and not a critical element of the early folding process. However, changes in the pathway need not significantly affect the overall shape of the free-energy profile as the effect of denaturant is expected to be largest along the reaction coordinate. Therefore, free-energy extrapolations may still yield good pictures of the reaction profile for folding, although one must always take into account variations in the ensemble composition at each point of the reaction coordinate.

2.6. Broad Barriers Mean Folding at Transition-State Level

Although the free-energy profiles presented here are purely phenomenological, they imply some new and interesting features of cooperativity and two-state folding. Broad barriers suggest that the denatured ensemble destabilizes instantly upon ordering and that folding propagates more or less isoenergetically at the transition-state level until the protein finally falls into the potential well of the native conformation. Disregarding the water, this may be visualized as an entropic uphill followed by a plateau at high free energy, where further entropic losses are precisely balanced by gains in intramolecular interactions, and eventually an enthalpic drop as the final contacts close within the nativelike but somewhat expanded structure. The folding events would then be determined by the sequence's ability to stabilize a productive set of transition states, rather than by restrictions in the denatured ground state.

3. Conclusions

Multistate kinetics and accumulation of compact denatured structures in protein folding are not always due to metastable intermediates but may be explained more simply by transient aggregation and changes of the protein folding transition state as a result of broad activation barriers. The findings reveal new features of the cooperativity in protein folding and unify a wide range of kinetic behaviors which have previously been taken as evidence for different folding mechanisms.

I wish to thank Alexei V. Finkelstein, Robert L. Baldwin, Peter G. Wolynes, Håkan Wennerström, Volodya Uverskii, and Daniel Otzen for valuable discussions and comments on the manuscript and the Swedish Natural Science Research Council, Crafordska Stiftelsen and S & E-C Hagbergs Stiftelse for support.

References

(1) Tanford, C. *Adv. Protein Chem.* **1968**, *23*, 121.
(2) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1990**, *59*, 631.
(3) Matouschek, A.; Kellis, J. T., Jr.; Serrano, L.; Bycroft, M.; Fersht, A. R. *Nature* **1990**, *346*, 440.
(4) Privalov, P. L. *J. Mol. Biol.* **1996**, *258*, 707.
(5) Clarke, A. R.; Waltho, J. P. *Curr. Opin. Biotechnol.* **1997**, *8*, 400.
(6) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1982**, *51*, 459.
(7) Creighton, T. E. *Curr. Biol.* **1991**, *1*, 8.
(8) Ikai, A.; Tanford, C. *Nature* **1971**, *230*, 100.
(9) Tsong, T. Y.; Baldwin, R. L.; Elson, E. L. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 2712.
(10) Baldwin, R. L. *J. Biomol. NMR* **1995**, *5*, 103.
(11) Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10869.
(12) Silow, M.; Oliveberg, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6084.
(13) Levinthal, C. *J. Chim. Phys.* **1968**, *85*, 44.
(14) Bryngelson, J. D.; Wolynes, P. G. *J. Chem. Phys.* **1989**, *93*, 6902.
(15) Sali, A.; Shakhnovich, E. I.; Karplus, M. *Nature* **1994**, *369*, 248.
(16) Bryngelson, J.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167.
(17) Finkelstein, A. V.; Badredtinov, A. Y. *Folding Des.* **1997**, *2*, 115.
(18) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10.
(19) Silow, M.; Oliveberg, M. *Biochemistry* **1997**, *36*, 7633.
(20) Roughton, F. J. W.; Chance, B. In *Techniques of Organic Chemistry*; Freiss, S. L., Lewis, E. S., Weissberger, A., Eds.; Wiley-Interscience: New York, 1963; pp 703−798.
(21) Baldwin, R. L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 84.
(22) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795.
(23) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1991**, *30*, 10428.
(24) Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266.
(25) Matouschek, A.; Serrano, L.; Fersht, A. R. *J. Mol. Biol.* **1992**, *224*, 819.
(26) Ewbank, J. J.; Creighton, T. E. *Biochemistry* **1993**, *32*, 3677.
(27) Pecorari, F.; Minard, P.; Desmadril, M.; Yon, J. M. *J. Biol. Chem.* **1996**, *271*, 5270.
(28) Oliveberg, M.; Tan, Y.-J.; Silow, M.; Fersht, A. R. *J. Mol. Biol.* **1998**, *277*, 933.
(29) Bernasconi, C. F. *Relaxation Kinetics*; Academic Press: New York, 1976.
(30) Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9029.
(31) Gutfreund, H. *Kinetics for the life sciences*; Cambride University Press: Oxford, U.K., 1995.
(32) Milla, M. E.; Sauer, R. T. *Biochemistry* **1994**, *33*, 1125.

(33) Tan, Y.-J.; Oliveberg, M.; Davis, B.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 980.
(34) Kiefhaber, T.; Rudolph, R.; Kohler, H. H.; Buchner, J. *Biotechnology* **1991**, *9*, 825.
(35) Cleland, J. L.; Wang, D. I. *Biotechnol. Prog.* **1992**, *8*, 97.
(36) Burton, R. E.; Huang, G. S.; Daugherty, M. A.; Fullbright, P. W.; Oas, T. G. *J. Mol. Biol.* **1996**, *263*, 311.
(37) Frauenfelder, H.; Wolynes, P. G. *Science* **1985**, *229*, 337.
(38) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. *Folding Des.* **1996**, *1*, 441.
(39) Plotkin, S. S.; Wang, J.; Wolynes, P. G. *J. Chem. Phys.* **1997**, *106*, 2932.
(40) Zwanzig, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 148.
(41) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11615.
(42) Matouschek, A.; Kellis, J. T., Jr.; Serrano, L.; Fersht, A. R. *Nature* **1989**, *342*, 122.
(43) Itzhaki, L. S.; Daniel, E. O.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 260−288.
(44) Lopez-Hernandes, I.; Serrano, L. *Folding Des.* **1996**, *2*, 43.
(45) Viguera, A. R.; Serrano, L.; Wilmanns, M. *Nat. Struct. Biol.* **1996**, *3*, 874.
(46) Parker, M. J.; Spencer, J.; Clarke, A. R. *J. Mol. Biol.* **1995**, *253*, 771.
(47) Tan, Y. J.; Oliveberg, M.; Fersht, A. R. *J. Mol. Biol.* **1996**, *264*, 377.
(48) Jonsson, T.; Waldburger, C. D.; Sauer, T. *Biochemistry* **1996**, *35*, 4795.
(49) Wolynes, P. G.; Luthey-Schulten, Z.; Onuchic, J. N. *Chem. Biol.* **1996**, *3*, 425.
(50) Matouschek, A.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7814.
(51) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334.
(52) Matouschek, A.; Matthews, J. M.; Johnson, C. M.; Fersht, A. R. *Protein Eng.* **1994**, *7*, 1089.
(53) Milla, M. E.; Brown, B. M.; Waldburger, C. D.; Sauer, R. T. *Biochemistry* **1995**, *34*, 13914.
(54) Chan, C. K.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1779.
(55) Matthews, J. M.; Fersht, A. R. *Biochemistry* **1995**, *34*, 6805.
(56) Creighton, T. Personal communication.
(57) Oliveberg, M. Unpublished results.
(58) Camacho, C. J.; Thirumalai, D. *Protein Sci.* **1996**, *5*, 1826.
(59) Zaidi, F. N.; Nath, U.; Udgaonkar, J. B. *Nature Struct. Biol.* **1997**, *4*, 1116.

AR970089M