UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.961 - 20

WORKSHOP ON:
PROTEINS, MEMBRANES and their INTERACTIONS

22 JULY - 2 AUGUST 1996

"Fast folding experiments and the topography of protein folding landscapes"

PART III

Zaida LUTHEY-SCHULTEN
University of Illinois
School of Chemical Science
505 South Mathews Avenue
IL 61801 Urbana
U.S.A.

Fast-folding experiments and the topography of protein folding energy landscapes

PG Wolynes[1], Z Luthey-Schulten[1] and JN Onuchic[2]

The rapid folding of certain proteins can be described theoretically using an energy landscape in the shape of a folding funnel. New techniques have allowed the examination of fast-folding events that occur in microseconds or less, and have tested the predictions of the theoretical models.

Addresses: [1]School of Chemical Sciences, University of Illinois, Urbana, IL 61801, USA and [2]Department of Physics, University of California, San Diego, La Jolla, CA 92093, USA.

Chemistry & Biology June 1996, **3**:425–432

 ISSN 1074-5521

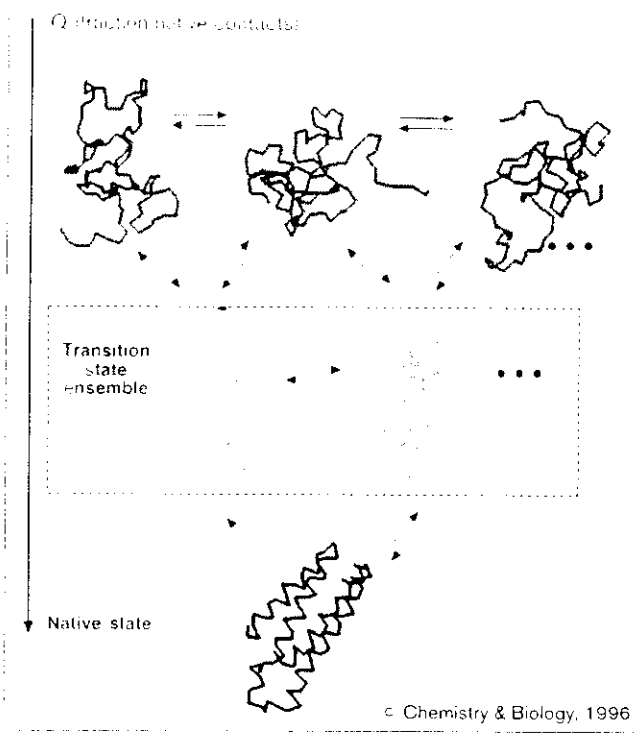
Introduction

For those who appreciate the beauty of the architecture of folded proteins it is amazing to find that such complex structures can be formed spontaneously in biologically short times, although doing so is clearly an evolutionary necessity. In human affairs, complex organized structures have always been made by a series of individually simple steps. A similar picture of the folding process seemed to emerge from the classical kinetic studies of folding on time scales greater than a few milliseconds [1]. But it was realized that, particularly for smaller proteins, folding could occur far more rapidly and without a requirement for this discrete stepwise assembly. This led to a view of protein folding based on a statistical mechanical characterization of the energy landscape of a folding protein ([2] and references therein), and this viewpoint has received considerable support from computational studies of model proteins [3–8]. With this new perspective it becomes clear that we will only be able to directly unravel the basic mechanisms of biomolecular self-organization and measure the topography of the folding-energy landscape with a new generation of experiments using relatively newer techniques that are capable of monitoring protein folding on shorter time scales (lasers [9–13], NMR [14–19], ultrafast mixing [20] and protein engineering [21,22]). This review highlights a remarkable fast folding experiment described in this issue by Mines *et al.* [23], who show experimentally a regularity in the folding of two proteins that is expected from a statistical description based on the landscape theory. We first discuss the theoretical descriptions of rapid protein folding before exploring to what extent the experiments of Mines *et al.*, and the increasing number of other recent fast folding experiments, support this description. Finally, we consider what the two fields have taught us so far about folding and the energy landscape.

The folding funnel

The energy-landscape theory of folding starts with the view that folding kinetics are best considered as a progressive organization of an ensemble of partially folded structures (Fig. 1), rather than a serial progression between intermediates. Thus, to understand kinetics one must statistically characterize the number and the distribution of free energies of the members of an ensemble of protein molecules that have partial order. A few collective reaction coordinates or order parameters should suffice to organize the description of the energy landscape. In contrast to most organic chemistry reactions, where the reaction coordinates describe a single structure, here they are collective coordinates characterizing an ensemble of structures as in electron transfer [24,25]. The search through these very large

Figure 1

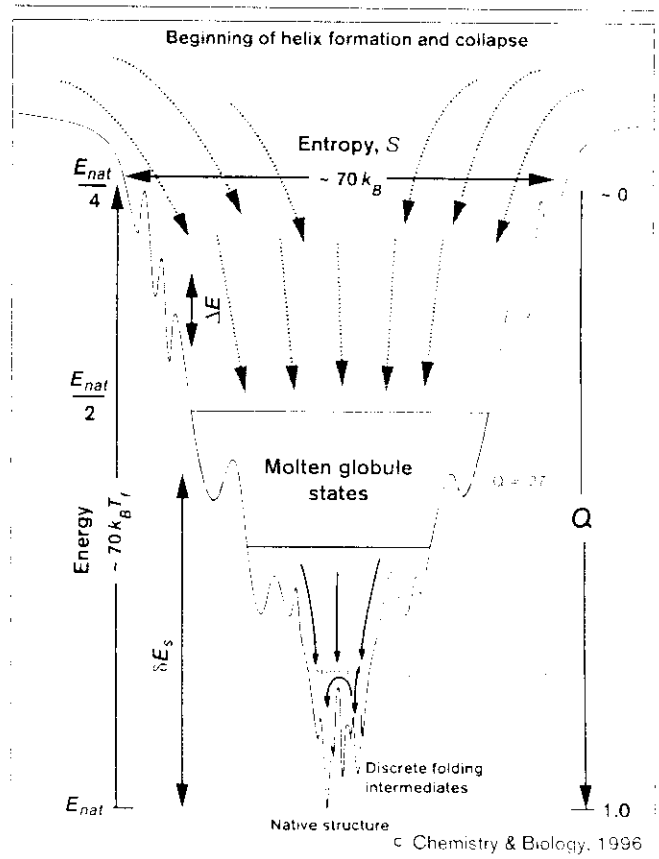


Network of conformational changes occurring in the ensemble of conformations for a protein as it folds to the native structure. The folding reaction coordinate Q is the average fraction of native contacts in members of the corresponding ensembles. The conformations in the transition state ensemble have a Q-value of ~0.6. The intermediate structures are obtained from computer simulations of a four helix bundle with varying degrees of native character. By organizing the states by energy and Q instead of the specific intermediates, a funnel description of the folding process shown in Figure 2 is obtained. (The simulations shown here are based on associative memory Hamiltonian structure prediction algorithms and show diminished helicity at low Q.)

number of configurations may present difficulties as pointed out by Levinthal years ago [26,27]. Even the protein conformations with given amounts of partial order have various energies that must be described by a distribution [28]. This give rise to what is called a 'rugged energy landscape', because many non-native interactions can exist in a (partially) disordered protein structure and these can be either favorable or non-favorable energetically. Conformational transitions from the deeper energy states in an ensemble must typically surmount an energy barrier determined both by this energetic ruggedness, ΔE[2], and any barriers provided by stereochemical constraints inherent in the polypeptide backbone.

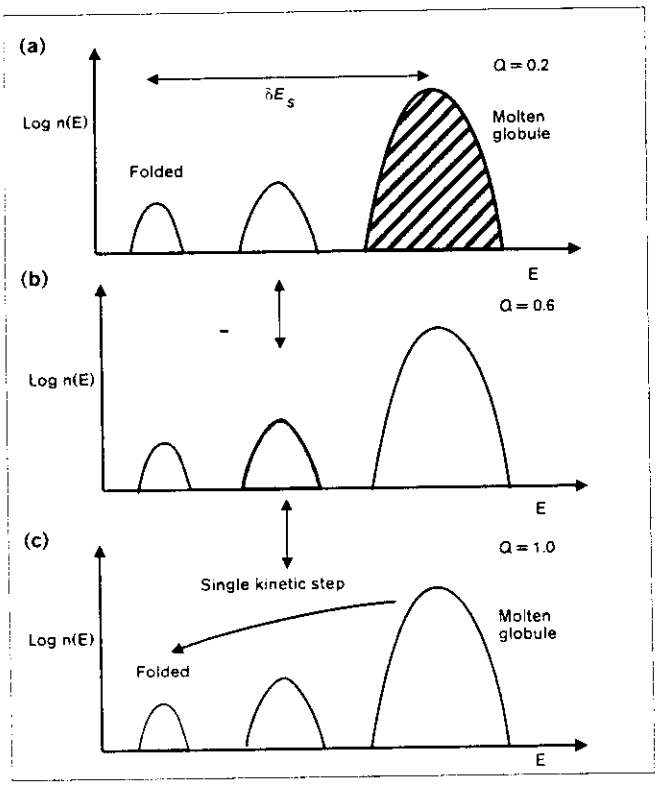
The ruggedness of the landscape is reflected in the kinetics of escape from transient conformational traps. Fast folding would be impossible on an entirely rugged landscape, which is called a completely frustrated landscape. Fast folding is only possible because of guiding forces that stabilize correct (native) interactions, on average, more than one expects by chance. We thus say that fast-folding proteins satisfy the principle of minimal frustration [28]. On a global level, the landscape must resemble a funnel ([29–31] and J.N.O., P.G.W., Z.L.-S. & N.D. Socci, unpublished data) (Fig. 2). As a folding reaction coordinate approaches its native value, the peak in energy distribution of the ensemble of conformations shifts to a lower value, as shown in Figure 3. The mean energy of structures in the funnel goes down as the ensemble approaches the native state. Simultaneously the number of conformations decreases, reflecting a loss of configurational entropy. (The entropy of the solvent, so important for the hydrophobic force, is included in the solvent average (free) energies used in the statistical distributions characterizing the landscape.) The

Figure 2



The schematic funnel for a 60-amino-acid helical protein as obtained from the corresponding-states principle analysis [30] with the 27-mer lattice model. The positions of the molten globule states, the transition state ensemble, and the local glass transition (the point at which discrete trapping states emerge) are shown as a function of the order parameters: fraction of native contacts Q, the solvents averaged energy E, and the configurational entropy S, which are drawn to scale. The stability gap δE_s quantitates the specificity of the native contacts. The ruggedness of the landscape ΔE determines the local barriers, but its value is not the conformational diffusion activation height, which is $E_a \sim 7k_BT_f$. k_B is the Boltzmann constant, T_f is the folding temperature, and E_{nat} is the energy of the native structure.

Figure 3

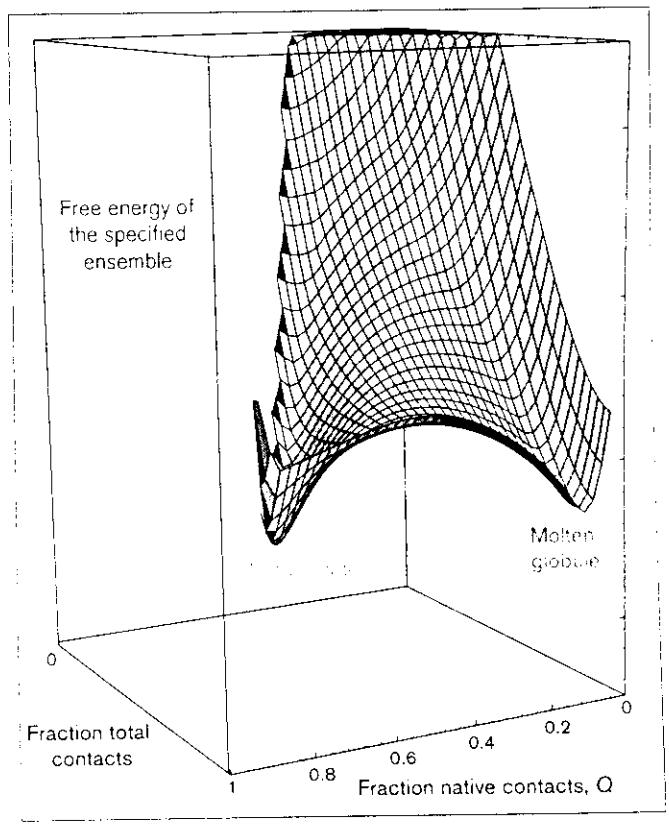


Histograms of protein configurational states at various stages of folding. **(a)** Collapsed phase with $Q = 0.2$. **(b)** Transition state with $Q = 0.6$ near the middle of the funnel. **(c)** Native state with $Q = 1.0$. In any given folding event, the protein molecule passes through intermediate Q states. In a type I folding, however, no significant population is trapped in these states, which are never substantially occupied kinetically. For downhill folding on fast time scales, these intermediate states may be sufficiently populated to be observed directly in the new fast-folding experiments. The y-axis is the logarithm of the number of configurations, and the x-axis is the free energy of a particular configuration averaged over the solvent.

mean energy loss down the funnel gives an additional energy parameter characterizing the landscape topography, the stability gap δE_s. The configurational entropy loss, the third significant parameter, is called the Levinthal entropy (S_f) as it is simply a logarithmic measure of the vastness of the configuration space of the biomolecule [30,32]. The fact that the landscape can be described by these very few parameters (δE_s, S_f, and ΔE) allows us to use a law of corresponding states (which uses scaling of certain known parameters to directly compare two different proteins) to understand the important characteristics of the funnel landscape for real proteins by investigating analytical models and lattice simulations [30]. This correspondence, and the quantitative one-dimensional funnel picture shown in Figure 2, are particularly easy to set up when collapse and non-specific secondary structure formation precede the rate-limiting stages of folding. A multidimensional funnel picture that includes additional energy scales for other collective reaction coordinates such as degree of collapse and local secondary structure has been recently developed (Z.L.-S., B.E. Ramirez & P.G.W., unpublished data) (Fig. 4). In this multidimensional picture the basic scenarios for folding remain the same, although new characteristic energies such as entropy diminution and energy loss upon collapse are needed to set up the correspondence. Fast folding experiments will allow the characterization of these important parameters, which determine a major part of the entropy loss, ultimately leading to folding.

The dominant dependence of folding on the stability gap (the difference in stability between folded and unfolded states) is strongly supported by the experiments of Gray, Winkler and colleagues [12,23]. Global changes to the

Figure 4



The multidimensional aspect of the free-energy surface. The average free energy of the ensemble (as opposed to that of an individual structure as shown in Fig. 2) is shown here as a function of the fraction of native contacts, Q, and total contacts. The free energies include the configurational entropy counting the number of states with a particular value of the collective coordinates. The surface is plotted using the analytical fit of Plotkin *et al.* [38], but the specific landmark values of Q are taken from the simulation results. At T_f we see two distinct minimal free energy ensembles: one (red) specifying the molten globule, the other (blue) specifying folded native-like configurations. The same transition-state ensemble (green) as in Figure 1 is represented in this lower dimensional view. The height of the thermodynamic barrier is approximately $2k_BT_f \approx 1.2$ kcal mol^{-1}.

molecule such as extensive sequence and solvent changes that lead to the same stability give a similar shaped funnel. If the corresponding-states approach were not valid, simulations of minimalist models could only be used as suggestive analogies to laboratory protein folding. Thus, this experiment encourages the application of landscape theory and simulations as quantitative tools.

Depending on the quantitative variation of entropy, mean energy and ruggedness as the protein ensemble descends in the funnel, energy-landscape theory provides a classification of kinetic scenarios [2]. If the energy loss always exceeds the entropy, we have a type 0 or 'downhill' folding scenario. Downhill folding on a relatively smooth landscape would resemble the physical process of polymer collapse and would not be expected to follow the simple exponential kinetics of a first order chemical reaction. For rougher landscapes, downhill processes would have an even wider distribution of time scales because of the varying depths of traps. These signatures should be quite obvious in fast folding experiments. For the roughest landscapes, only a few discrete traps could be populated and the dynamics could be described by a discrete but complex kinetic scheme, after the resolution of the early events. A different result is seen when the rate of energy loss down the funnel does not match, at each stage, the entropy loss as the ensemble descends. As the free energy is defined as $E - TS$, a free energy maximum will result when the entropy decreases more rapidly than the energy. In this type I scenario, a bottleneck occurs in the folding process. The ensemble of structures at the bottleneck represents a free energy barrier, and at the macroscopic level the folding can be described using single-step, exponential kinetics. Landscape theory predicts simple regularities for the folding-rate coefficient when the type I scenario applies. As a large number of different configurations are present at the bottleneck, detailed energetic effects on single configurations get averaged out, and thus the folding rate is largely and smoothly related to the stability.

Transfer coefficients (Φ) in local linear free energy relations (LFER) reflect the location of the bottleneck ensemble along the reaction coordinate. Using the assumption that bottleneck position does not change much when the landscape is changed, we can derive an LFER between the folding time, τ_f, and the free-energy change in folding, which to the same order is $< \Delta H >_u - < \Delta H >_f$ (J.N.O., P.G.W., Z.L.-S. & N.D. Socci, unpublished data), where the brackets indicate averaging over an ensemble. Thus:

$$\frac{\partial \log \tau_f}{\partial \log K} = \Phi = \frac{< \Delta H >_{ts} - < \Delta H >_u}{< \Delta H >_u - < \Delta H >_f}$$

where K is the equilibrium constant for the folding reaction and the subscripts u, f and ts indicate the unfolded, folded and transition-state ensembles, respectively. For global perturbations, Φ gives a good first approximation of the position of this bottleneck. Φ will be strongly correlated to the degree of nativeness of the transition-state ensemble. Only an average description is obtained, however, using a single value of Φ. Proteins are polymeric leading to varying participation of different residues in the transition-state ensemble, and protein engineering techniques reveal a histogram of participation of residues in the configurations of a transition state ensemble, centered around a fractional value ([21] and J.N.O., P.G.W., Z.L.-S. & N.D. Socci, unpublished data). Of course, as increased stabilization can transform a type I scenario to a downhill one, some curvature of the LFER is also to be expected.

Type II scenarios occur if the ruggedness of the landscape is very large, so that only a few discrete trap configurations can be thermally occupied, even before the bottleneck is kinetically surmounted (type IIB). This transition to discrete long-lived traps in the type IIB scenario has much in common with the glass transition, where disordered configurations of a liquid can exist for long periods of time. Although the rough location of transition to this regime can be inferred from simple statistical mechanical theories involving the interplay of ruggedness and entropy, the most important result of landscape theory for the kineticist is the prediction that the existence of each specific trap is very sensitive to even single-site changes in the molecule and might be eliminated by simple protein engineering or, sometimes, by chemistry [17,33]. Many of the classically studied folding processes encountered type II scenarios either after a downhill run or after encountering an especially small entropic barrier at the bottleneck ensemble. Landscape theory for type II folding can only be tested quantitatively with combinatorial experiments testing the distribution of intermediate lifetimes using a variety of different sequences. The fast folding events, on the other hand, should conform to type 0 or type I scenarios and therefore will exhibit greater regularities and be dominated by the guiding forces for folding rather than ruggedness and trapping.

Fast folding experiments to date

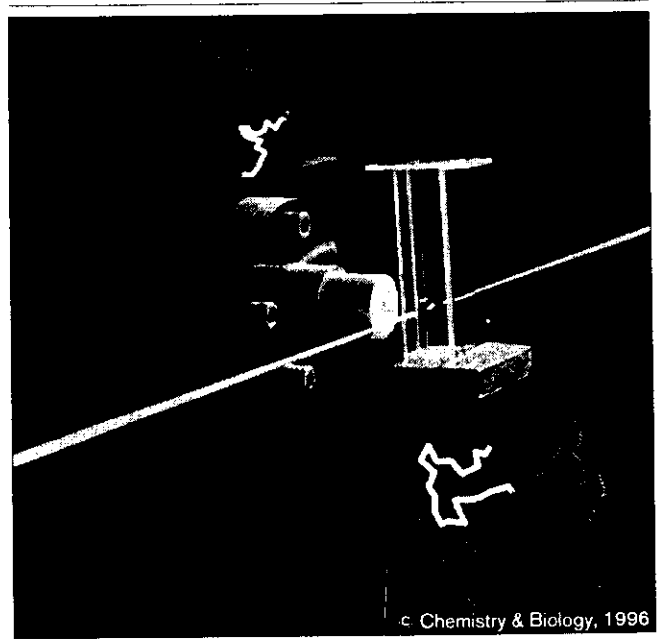
To initiate fast folding, the protein in question must somehow be converted, almost instantaneously, from a form that favors the unfolded state to a form that favors the folded state. Several methods have recently been developed to trigger this event. The first protein to be explicitly studied on the fast (submillisecond) time scale was cytochrome *c*, in an effort led by Eaton and Roder [9]. Previously, a burst phase of rapid apparent folding had been observed in stopped flow, in which signatures of native structure were seen within the dead-time of the apparatus [17]. It was also shown that the slow phases of folding were largely eliminated under pH conditions where a specific histidine is protonated and thus unable

to misassociate with the heme iron, leading to a kinetic trap [33]. Jones *et al.* [9] initiated folding photochemically by exploiting the fact that the stability of the protein is enhanced by coordination with two sidechains of the polypeptide chain. Cytochrome *c* with bound carbon monoxide thus denatures more readily than the protein that lacks CO, and flash photolysis of the CO-bound form with a laser can initiate folding within picoseconds under appropriate denaturing conditions. The change in stability is modest, so the experiment must be performed with a relatively high denaturant concentration to ensure that the initial CO-bound form is completely denatured. This procedure leads to type I folding in about 10 ms. Significantly, Jones *et al.* also observe complex spectral changes, beginning in the microsecond range. These relaxation events occur exclusively in the unfolded phase, but unfortunately all the free energy gradient in this case is caused by the differing chemical stability of the amino acids in the chain that can act as ligands of the heme and not by the dominant folding forces. Nevertheless combining these measurements of the dynamics in the unfolded state with a separate determination of the inherent chemical effects on ligand association has led to a plausible upper limit on the speed of downhill protein folding of 1 μs [34].

Gray and colleagues [12,23] exploit a conceptually similar route to fast folding. Reduced and oxidized cytochromes have different stabilities, and at a particular concentration of denaturant, the oxidized protein is unfolded and the reduced protein is folded. Thus photo-initiated electron transfer can induce folding. The oxidized, unfolded protein is bombarded with electrons, and upon conversion to the reduced form, folding is instantly favored. Using a photosensitizer electron injection system, they were able to initiate folding in less than 1 μs [12]. As the reduced cytochrome *c* in this case had a lifetime of only ~1 ms, a different injection system was used to study longer periods of time. Although it causes a larger stability change than CO binding, the electron transfer initiated folding was still studied under type I scenario conditions, giving results consistent with the earlier work, while reducing some technical difficulties arising from back reactions.

The experiment described in this issue by Mines *et al.* [23] also uses electron-transfer triggering to show a remarkable regularity that is expected for type I transitions. The part of the folding funnel probed in a type I scenario is predicted to be self-averaging, so two quite different sequences folding to the same structure should have funnel shapes in the region of high density of states (above the transition states) that depend mostly on the overall stability and amino-acid composition of the molecule. The experiments of Mines *et al.* show that horse and yeast cytochrome *c* have the same folding rates when tuned to the same stability by adjusting the denaturant

Figure 5



A composite picture is shown of the sample cell and laser T-jump apparatus of Ballew *et al.* [13] that was used to measure the folding of apomyoglobin from its cold-denatured state and the computer simulations of two conformations with a low (top) and high (bottom) fraction of native contacts. The conformations were obtained from a protein structure prediction simulation using an associative memory energy function [39] and are color-coded to indicate the foldon regions [35] that could be kinetically competent quasi-independent folding units.

concentration. This is true even though the sequences are only 50 % identical. This non-sequence-specificity of type I funnel-like folding contrasts greatly with the dynamics of type II folding; for example, the late-stage folding intermediates of lactalbumin and hen lysozyme are quite different [18]. This is true despite the fact that the two proteins share 36 % sequence identity.

Laser fast folding experiments can examine the dynamics of a larger portion of the folding events only by obtaining bigger stability changes. This can be achieved with a T-jump. The record for initiating folding thus far has been achieved in groundbreaking experiments of Gruebele and colleagues [13], schematically presented in Figure 5. Laser heating using an infrared pulse allows a protein that is denatured because of the cold to be taken to conditions that favor the native form within nanoseconds.

This technique to has been applied to the study of apomyoglobin, which is denatured by cold at a convenient temperature near the freezing point of water. Conformational changes can then be monitored by measuring fluorescence decay of the tryptophan by repetitive femtosecond pulses after initiating folding. Modeling the fluorescence decays at equilibrium leads to the conclusion that

the change that is detected between native and non-native states is primarily determined by a specific contact between a methionine and the tryptophan. Thus the experiment measures part of the tertiary structural transition. The data can be fit by a largely exponential process, indicating a type I scenario for this conformational change, which occurs in 7 μs. There are also some signs of shorter time scale changes in the decay, which may be connected with secondary structure formation or more local collapse of the polypeptide chain. The assignment of the major spectral change to a large-scale conformational movement is suggested by its viscosity dependence, as shown by adding glycerol to the solution. This study confirms a Kramers-like dependence of the rate on external viscosity [31]. Longer time scale measurements show that the transition seen by Gruebele and coworkers is not complete folding [19]. At least in sperm whale apomyoglobin, the complete folding occurs later. This may suggest the presence of a subdomain structure in the apomyoglobin molecule or specific traps during folding [35].

The observations of interesting dynamics on the microsecond timescale with cytochrome *c* and apomyoglobin are also confirmed in a fast T-jump experiment on barstar (barnase inhibitor) [36]. A rapid transition taking 300 μs is followed by a 100-ms slow phase. The rapid change involves a very small amplitude of fluorescence change but is clearly not an artifact. As in the experiments with cytochrome *c* [23], the location of the transition state (Φ) for this folding is early along the folding reaction coordinate as judged by its denaturant dependence. Clearly many interesting folding processes occur on the microsecond to millisecond timescale. These processes can also be observed using an enhancement of the conventional mixing techniques using small orifices. One such ultrafast mixing experiment with cytochrome *c* shows a folding process that is strongly extended and nonexponential, suggesting the possibility of a downhill or type 0 folding scenario [20]. This novel technique is still in its early stages and several complexities concerning the interaction of the turbulent hydrodynamic flows and the polymer conformations need to be sorted out in order to be completely confident of the results. Notice here that there is an apparent change of the Φ values, suggesting that nonlinearity of the free energy relation can be observed when experiments are done over a large enough range of denaturant concentrations.

NMR spectral line shape analysis can be used, at least near the midpoint of the folding curve, to study submillisecond folding dynamics. This technique has been used to examine both folding and unfolding of monomeric λ repressor [14]. This molecule is highly helical and closely resembles many of the systems studied by the theoreticians. Again an intermediate Φ value is observed showing the existence of a bottleneck or a transition state ensemble in a type I scenario near to the location predicted from theory.

Although most of the action in fast folding seems to occur in the range of microseconds, very fast processes have also been observed in unfolding experiments, such as those examining the thermally induced unfolding of apomyoglobin [11]. Protein conformation is probed using vibrational spectroscopy techniques that are sensitive to secondary structural changes. Interestingly, if the process is two-state, one of the fast unfolding processes observed could be interpreted to give a refolding time constant close to that determined by laser heating, if the temperature dependence of viscosity is taken into account. Transient infrared spectroscopy has also been used by monitor events occurring in nanosecond timescales [10]. These experiments show fairly complicated kinetics that cannot be fit by a single-step representation. Some changes occur on such short timescales that they may be accessible to direct molecular dynamics simulation [37]. Clearly a rich variety of phenomenology is beginning to appear from the submillisecond folding experiments. While many interpretational questions remain to be answered, it is quite clear that events that were earlier postulated based only on slow stopped-flow mixing experiments were just a small part of what is going on.

Quantifying funnel shapes with fast-folding experiments

We are just beginning to understand protein folding quantitatively in terms of energy-landscape topography. The fast-folding experiments we have just discussed support the qualitative notions of the theory and begin to allow us to refine the quantitative aspects of the funnel shape. Foremost among the issues in funnel shape are the questions of the appropriate coordinates for describing the crucial, partially folded ensembles of structures. In the simplest model, the fraction of native contacts (Q) is used. Because each experiment has so far used only a small fraction of the possible probes, this very basic issue still requires much study. Clearly the important configurations, as revealed in these experiments, are partially collapsed and probably have some secondary structure formed, and even though parts of the secondary structure may be in the correct locations, only a small amount of tertiary ordering may exist. Many of the next generation of experiments will begin to address these issues more completely using other techniques. Circular dichroism would be a more convincing measure of overall secondary structure, while isotopic labeling and vibrational spectroscopy could locate the precise secondary structural elements [11].

More refined fluorescence experiments using mutants with tryptophan residues placed in varying locations could pinpoint tertiary contacts. Even without the precise definition of the relevant coordinates, some quantitative features — the entropic aspects of the landscape and the ruggedness — are grossly revealed by the experiments to date and seem to have a reasonable relationship to the theoretical expectations. The fast folding experiments all suggest that

the bottleneck for folding at the midpoint of the denaturing curve corresponds to a transition state ensemble with a considerable configurational entropy. The fractional degree of folding as determined from linear free-energy relationships agrees reasonably well with results obtained from lattice simulations as expected from the corresponding states principle interpretation of the funnel landscape. This is probably the best current measure of the entropic effects in the funnel. More detailed studies of the specific contacts formed in the transition state ensemble are also in agreement with the picture that a collection of delocalized folding nuclei are important.

The ruggedness of the landscape can be quantified more accurately using the temperature dependence of the folding rate when under type I conditions or by the detailed time course in a type 0 downhill folding scenario. The data from the cytochrome c experiments show that configurational diffusion is faster at high temperatures than at low, as the isostability values of the folding rate are higher at 40 °C than at 22.5 °C [12,23]. Some of this effect can be attributed to the change in viscosity with temperature, as would be expected from the experiments with apomyoglobin [13], but this would only account for a small fraction of the effect. The ratios of the viscosity at these two temperatures is ~1.5, but the rate ratio is ~20. We are only able to calculate a value for the apparent activation barrier by assuming that the entire ruggedness is enthalpic (whereas it is doubtless partially entropic due to the hydrophobic effects); nevertheless, with this imperfect description we arrive at a value of $E_a \sim 11k_BT_f$. We should be very careful when providing an interpretation of such an apparent barrier [27,31]. The simplest description, using the configurational diffusion model of Bryngelson and Wolynes [27], directly relates this barrier to the average roughness of the landscape. According to their model, at temperatures sufficiently high compared to the glass transition temperature, $E_a \sim \Delta E^2/2k_BT_f \approx 10k_BT_f$. Below the glass transition temperature, the system becomes caught in a few long-lived non-native states, making the average ruggedness description invalid. The experimental value obtained for this apparent barrier is slightly larger than the actual simulated value, $E_a \sim 7k_BT_f$, using the corresponding-state analysis for a 60 residue protein. Estimates of the ruggedness and conformational diffusion activation barriers can be refined by much more extensive studies in a larger temperature and viscosity range, both for the data of Mines *et al.* [23] with a type I scenario, and the more detailed time course of an apparently type 0 scenario observed in the ultrafast mixing experiment [20].

Another fact that emerges from the fast folding experiments is that there are important fast-folding substructures in the larger proteins. In any event, the simple one-parameter funnel picture that we have highlighted in the early part of this review should apply for the smaller protein systems. The experiments with apomyoglobin suggest that a unit consisting of the A, G and H helices forms first. This unit is similar to the size of proteins described by a single funnel but clearly the overall folding may require a funnel for each substructure. A theoretical description of folding in domains or foldons has already been made using energy-landscape analysis [35].

Fast folding experiments provide the most direct view on the forces that govern the self-organization of protein molecules. The quantification of energy-landscape topography that these experiments will make possible should help structure prediction schemes that are already based on landscape ideas, but which have been handicapped by only knowing the final folded structures of proteins. Thus, we can look forward to a very fruitful interaction between the experimental community and the computational community.

Acknowledgements

The work at the University of California at San Diego was supported by the NSF (Grant No. MCB-93-16186) and at Illinois by the NIH (Grant No. 2 R01 GM44557).

References

1. Kim, P.S. & Baldwin, R.L. (1990). Intermediates in the folding reactions of small proteins. Annu. Rev. Biochem. **59**, 631–660.
2. Bryngelson, J.D., Onuchic, J.N., Socci, N.D., & Wolynes, P.G. (1995). Funnels, pathways and the energy landscape of protein folding: a synthesis. Proteins **21**, 167–195.
3. Friedrichs, M. & Wolynes, P.G. (1990). Molecular dynamics of associative memory Hamiltonians for protein tertiary structure recognition. Tetrahedron (Comp. Meth.) **3**, 175–190.
4. Dill, K.A., et al. & Chan, H.S. (1995). Principles of protein folding: a perspective from simple exact models. Protein Sci. **4**, 561–602.
5. Honeycutt, J.D. & Thirumalai, D. (1992). The nature of the folded state of globular proteins. Biopolymers **32**, 695–709.
6. Sali, A., Shakhnovich, E. & Karplus, M. (1994). Kinetics of protein folding: a lattice model study of the requirements for folding to the native state. J. Mol. Biol. **235**, 1614–1636.
7. Socci, N.D. & Onuchic, J.N. (1994). Folding kinetics of protein-like heteropolymers. J. Chem. Phys. **101**, 1519–1528.
8. Socci, N.D. & Onuchic, J.N. (1995). Kinetic and thermodynamic analysis of protein-like heteropolymers: Monte Carlo histogram technique. J. Chem. Phys. **103**, 4732–4744.
9. Jones, C.M., et al. & Eaton, W.A. (1993). Fast events in protein folding initiated by nanosecond laser photolysis. Proc. Natl. Acad. Sci. USA **90**, 11860–11864.
10. Phillips, C.M., Mizutani, Y. & Hochstrasser, R.M. (1995). Ultrafast thermally induced unfolding of RNase A. Proc. Natl. Acad. Sci. USA **92**, 7292–7296.
11. Williams, S., et al. & Dyer, R.B. (1996). Fast events in protein folding, helix melting and formation in a small peptide. Biochemistry **36**, 691–698.
12. Pascher, T., Chesick, J.P., Winkler, J.R. & Gray, H.B. (1996). Protein folding triggered by electron transfer. Science **271**, 1558–1560.
13. Ballew, R.M., Sabelko, J. & Gruebele, M. (1996). Direct observation of fast protein folding: the initial collapse of apomyoglobin. Proc. Natl. Acad. USA **93**, 5759–5764.
14. Huang, G.S. & Oas, T.G. (1995). Submillisecond folding of monomeric λ repressor. Proc. Natl. Acad. Sci. **92**, 6878–6882.
15. Feng, Y., Sligar, S.G. & Wand, A.J. (1994). The solution structure of apocytochrome b_{562}. Nat. Struct. **1**, 30–36.
16. Nash, D., Lee, B. & Jonas, J. (1996). Hydrogen exchange kinetics in the cold denatured state of ribonuclease A. Biochim. Biophys., Acta in press.
17. Roder, H. & Elove, G.A. (1994). Early stages of protein folding. In

Mechanisms of Protein Folding (Pain, R.H., ed.). pp. 26–54, Oxford Press, Oxford.
18. Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J. & Dobson, C.M. (1993). Structure and dynamics of the acid-denatured molten globule state of α-lactalbumin: a two-dimensional NMR study. Biochemistry **32**, 1707–1718.
19. Wright, P.E. & Jennings, P.A. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. Science **262**, 892–895.
20. Chan, C-K., Hu, Y., Takahashi, S., Rousseau, D.L., Eaton, W.A. & Hofrichter, J. (1996). Protein folding kinetics studied by ultrafast mixing. Biophys. **70**, A177.
21. Itzhaki, L.S., Otzen, D.E. & Fersht, A.R. (1995). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. J. Mol. Biol. **254**, 260–288.
22. López-Hernández, E. & Serrano, L. (1996). Structure of the transition state for folding of the 129 amino acid protein CheY resembles that of a smaller protein, CI-2. Fold. Des. **1**, 43–55.
23. Mines, G.A., Pascher, T., Lee, S.C., Winkler, J.R. & Gray, H.B. (1996). Cytochrome c folding triggered by electron transfer. Chemistry & Biology **3**, 491–497.
24. Onuchic, J.N. & Wolynes, P.G. (1993). Energy landscapes, glass transitions and chemical reaction dynamics in biomolecular or solvent environment. J. Chem. Phys. **98**, 2218–2224.
25. Marcus, R.A. (1993). Electron transfer reactions in chemistry: theory and experiment. Rev. Mod. Phys. **65**, 599–610.
26. Levinthal, C. (1969). How to fold graciously. In Mossbauer Spectroscopy in Biological (DeBrunner, P., Tsibris, J. & Munck, E., eds), pp. 22–24, University of Illinois Press, Urbana, IL.
27. Bryngelson, J.D. & Wolynes, P.G. (1989). Intermediates and barrier crossing in a random energy model (with applications to protein folding). J. Phys. Chem. **93**, 6902–6915.
28. Bryngelson, J.D. & Wolynes, P.G. (1987). Spin glasses and the statistical mechanics of protein folding. Proc. Natl. Acad. Sci. USA **84**, 7524–7528.
29. Leopold, P.E., Montal, M. & Onuchic, J.N. (1992). Protein folding funnels: a kinetic approach to the sequence–structure relationship. Proc. Natl. Acad. Sci. USA **89**, 8721–8725.
30. Onuchic, J.N., Wolynes, P.G., Luthey-Schulten, Z. & Socci, N.D. (1995). Towards an outline of the topography of a realistic protein folding funnel. Proc. Natl. Acad. Sci. USA **92**, 3626–3630.
31. Socci, N.D., Onuchic, J.N. & Wolynes, P.G. (1996). Diffusive dynamics of the reaction coordinate for protein folding funnels. J. Chem. Phys. **104**, 5860–5868.
32. Luthey-Schulten, Z.A., Ramirez, B.E. & Wolynes, P.G. (1995). Helix–coil, liquid–crystal and spin-glass transitions of a collapsed heteropolymer. J. Phys. Chem. **99**, 2177–2185.
33. Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. (1994). The barriers in protein folding. Nat. Struct. Biol. **1**, 149–156.
34. Hagen, S.J., Hofrichter, J., Szabo, A. & Eaton, W.A. (1996). Diffusion limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. Proc. Natl. Acad. Sci. USA in press.
35. Panchenko, A.R., Luthey-Schulten, Z.A. & Wolynes, P.G. (1996). Foldons, protein structural modules and exons. Proc. Natl. Acad. Sci. USA **93**, 2008–2013.
36. Nolting, B., Golbik, R. & Fersht, A.R. (1995). Submillisecond events in protein folding. Proc. Natl. Acad. Sci. USA **92**, 10668–10672.
37. Daggett, V. & Levitt, M. (1992). A model of the molten globule state from molecular dynamics simulations. Proc. Natl. Acad. Sci. USA **89**, 5142–5146.
38. Plotkin, S.S., Wang, J. & Wolynes, P.G. (1996). A correlated energy landscape model for finite, random heteropolymers. Phys. Rev. in press.
39. Goldstein, R.A. Luthey-Schulten, Z.A. & Wolynes, P.G. (1992). Optimal protein-folding codes from spin-glass theory. Proc. Natl. Acad. Sci. USA **89**, 4918–4922.

Direct observation of fast protein folding: The initial collapse of apomyoglobin

(molten globule/temperature jump/fluorescence/circular dichroism)

R. M. BALLEW, J. SABELKO, AND M. GRUEBELE

School of Chemical Sciences and Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana, IL 61801

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 5759–5764, June 1996
Chemistry

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 5759–5764, June 1996
Chemistry

Direct observation of fast protein folding: The initial collapse of apomyoglobin

(molten globule/temperature jump/fluorescence/circular dichroism)

R. M. BALLEW, J. SABELKO, AND M. GRUEBELE

School of Chemical Sciences and Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana, IL 61801

Communicated by Peter G. Wolynes, University of Illinois, Urbana, IL, February 5, 1996 (received for review January 25, 1996)

ABSTRACT **The rapid refolding dynamics of apomyoglobin are followed by a new temperature-jump fluorescence technique on a 15-ns to 0.5-ms time scale *in vitro*. The apparatus measures the protein-folding history in a single sweep in standard aqueous buffers. The earliest steps during folding to a compact state are observed and are complete in under 20 μs. Experiments on mutants and consideration of steady-state CD and fluorescence spectra indicate that the observed microsecond phase monitors assembly of an A·(H·G) helix subunit. Measurements at different viscosities indicate diffusive behavior even at low viscosities, in agreement with motions of a solvent-exposed protein during the initial collapse.**

A compact molten globule state has been proposed as an early protein-folding intermediate in many cases (1), but its initial formation from the unfolded state during refolding has not been directly observed. Slower (>1-ms) kinetic phases of protein folding, leading from the molten globule state to the native state (2) or from the unfolded state to the fully native state directly (3, 4), have been resolved by stopped-flow techniques. The nanosecond to sub-millisecond time scale, in which the earliest global protein-folding motions are expected to occur, is not easily accessible by these techniques.

Recently, optically induced ligand unbinding in cytochrome *c* (5) and rapid mixing techniques (6) have been used to monitor folding dynamics on a sub-millisecond time scale. The helix-coil transition in peptides and unfolding of myoglobin (Mb) (7), and the unfolding of RNase A (8) have been monitored by infrared absorption after a temperature jump (T-jump). This very early regime has also been accessed by NMR line shape measurements (9) and can be reached by electron-transfer-induced folding (10). Finally, the microsecond T-jump technique developed by Eigen and coworkers (for review, see ref. 11) has been applied to folding experiments starting with cold-denatured barstar (12).

Theoretical models currently under discussion address various issues of early folding. A general principle that has emerged is that of "minimal frustration": a free-energy funnel to the native state provides enough (perhaps just enough) smoothness to balance the natural roughness of the folding free-energy surface expected for a heteropolymer, thus allowing fairly efficient folding (13). Depending on the relative location of the transition state, the glass-like transition, and the roughness and sloping of the free-energy surface, this scenario has been played out analytically and in simulations in many limits, including direct two-state folding, collapse to folding intermediates with varying amounts of secondary and tertiary structure, and kinetic traps (14–17). One certainty is that, at least transiently, the protein must pass through a compact ensemble of states while folding.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

To investigate early folding events in globular proteins, we have developed an apparatus that allows us to monitor the T-jump-induced refolding dynamics of small proteins. It extends time resolution to the nanosecond time scale and temperature differentials to 30 K, allowing easier observation of larger population changes at shorter times. Some of the requirements in developing our approach were as follows: (*i*) nanosecond to millisecond time coverage with nanosecond resolution and dead time to follow the earliest large-scale backbone motions up to the stopped-flow time regime; (*ii*) single-shot acquisition of a sample's history without pump-probe signal averaging or sample flow to allow for small sample quantities (e.g., genetically engineered ones); (*iii*) experiments in simple aqueous buffers, obviating the need for extraneous dyes or other molecules that could affect early folding dynamics; and (*iv*) the possibility of fluorescence, CD, or infrared monitoring in different viscosity, temperature, or denaturant concentration ranges.

EXPERIMENTAL METHODS

Outline. The heart of the experiment is shown in Fig. 1. A protein sample is cold denatured in a short path length cell by supercooling the aqueous buffer. CD is used to verify unfolding under the conditions used. The aqueous buffer is directly and rapidly heated (up to $5 \cdot 10^9$ K/s) by a nanosecond infrared Raman pulse. The unfolded protein is now located in a warmed buffer at a temperature thermodynamically conducive to folding. The folding process is followed by focusing a train of UV laser pulses onto the sample. Here, we report on fluorescence experiments using Trp excitation at 280 nm. Every 15 ns, the UV pulse train induces a protein fluorescence transient. The transients are collected continuously by a photomultiplier and transient digitizer. Finally, the changes in protein fluorescence in 15-ns or larger (averaged) steps are analyzed and correlated to structural changes.

Heating and Sample Cell. The sample is held in a custom 0.4-mm path length-fused silica cell (Fig. 1) cooled by two thermoelectric devices. The temperature is held constant to <0.2°C by a thermistor feedback loop. Two ≤120-mJ, 1.54-μm infrared beams are generated by Raman shifting a 700-mJ neodymium:yttrium/aluminum-garnet laser in a mode-optimized high-efficiency methane cell, resulting in uniform, near-gaussian heating profiles of 2-mm diameter. The counterpropagating beams are delayed by 8 ns from one another to avoid transient grating formation, and heating (by OH overtone relaxation in water) is completed within the pulse duration due to picosecond vibrational equilibration. The two mirror-image exponential absorption profiles add up to a longitudinal temperature uniformity of ±3% over the length of the cell, and the large uniform pump profile minimizes thermal lensing and diffusion effects. The T-jump is measured by transmission of a 1.5-μm diode laser focused to <400 μm

Abbreviations: T-jump, temperature jump; Mb, myoglobin; apoMb, apomyoglobin; h-apoMb, horse apoMb.

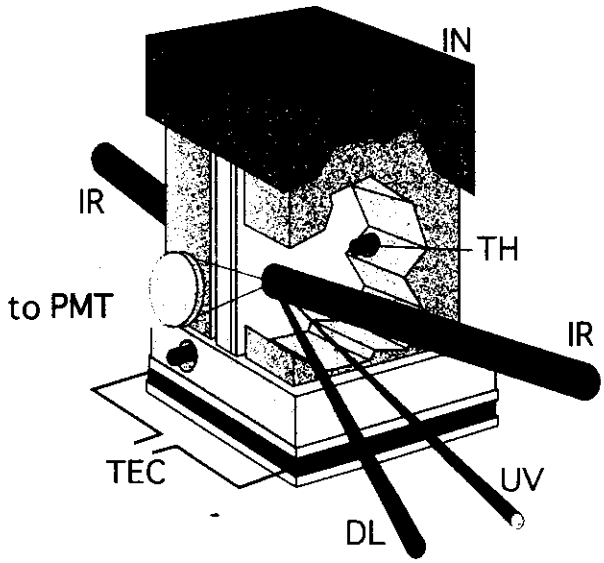


FIG. 1. Cell and beam geometry. Counterpropagating infrared beams (IR) heat the sample. The ultraviolet beam (UV) probes the fluorescence and a low-power diode laser (DL) monitors temperature by changes in transmission. The cell is encased in an aluminum block shielded by heavy insulation (IN) and cooled by a thermoelectric device (TEC) with a feedback loop. Several thermistors (TH) monitor the temperature below and at beam height. Trp fluorescence is collected at right angles by $f/0.8$ optics. The entire apparatus is enclosed in a dry nitrogen box to prevent condensation.

and calibrated at known temperatures or with free Trp lifetimes. The custom cell design reduces diffusional and shock-wave effects in the sample. The latter are also minimized by the 4°C density maximum of water. Any residual shock-wave effects and weak anti-Stokes Raman/supercontinuum flashes from the infrared optics do not affect lifetime and folding measurements, as discussed in *Fast Kinetic Results* (see also Fig. 6).

Kinetic Probing. Probe pulses for Trp fluorescence are generated by tripling a self-mode-locked titanium:sapphire laser in BBO nonlinear crystals. The focused pulse train (≈2 mW, 7-nm excitation bandwidth at 280 nm, and 200-μm diameter) probes only the center of the large pump profile. Fluorescence is imaged onto a filtered (U340) 600-ps rise-time micro-photomultiplier tube connected to a 2-GS/s digitizer (0.75-GHz bandwidth, 1 million channels). All infrared and UV probe beams are electromechanically shuttered during the experiment to avoid optical sample denaturation.

Protein Samples. Horse skeletal muscle Mb was obtained from Sigma. The heme was removed by the 2-butanone method (18). Purity of the apomyoglobin (apoMb) was verified by UV-visible spectroscopy to exceed 99%. Concentrations of apoMb were determined as described (19). All samples were buffered in 10 mM sodium acetate at pH values shown in Table 1. A Trp-7 → Phe-7 sperm whale Mb mutant was provided by S. Sligar (20) and converted to apoMb as described above.

Summary. The folding history of the protein in an aqueous buffer is thus obtained in a single shot in real time, requiring no dyes for energy transfer, extensive pump-probe signal averaging, or large samples for sample flow. The dependence of the early folding kinetics on the protein environment can be studied by changing solvent viscosity (e.g., glycerol solutions) or by adding denaturants.

STEADY-STATE RESULTS

We have chosen horse apoMb (h-apoMb) for our first experiment because a wealth of information on its cold denaturation (21), CD (22), millisecond folding dynamics (2), structure (23), molten globule states (24), and fluorescence properties (25) is available. Our cold denaturation data on h-apoMb are similar to the results obtained by Nishii and coworkers (22). Assuming a two-state model with folded/unfolded states, the population of the native state as a function of temperature is plotted in Fig. 2. The cold-denatured state probably contains some residual helix, although the heat- and cold-denatured baseline in Fig. 2 is well-fitted by a single line. Cold and pressure denaturation experiments on RNase by Nash and coworkers (26) indicate protection factors for a few residues in the pressure/cold-denatured state that are substantially higher (≈100) than those in the heat-denatured state (≈1–5) but substantially lower than those in the native protein (10^3–10^4), and the higher cold denaturation transition temperature of apoMb implies even smaller protection factors for the latter. Calorimetric measurements also suggest a state at least close to a random coil (21), and a number of steady-state and kinetic fluorescence experiments described below also indicate a loose initial state. However, until isotopic two-dimensional NMR or similar data become available, a partially assembled A·G·H complex cannot be ruled out entirely in the cold-denatured initial state. It is worth reiterating that our experiment does not take place under cold denaturation conditions; the conditions simply provide an initial unfolded structure from which the protein starts refolding in the warmed buffer.

Fluorescence/mutagenesis experiments have shown that specific amino acid residues can dominate Trp quenching (27, 28). In certain cases, this is the cause for observed nonexponential behavior in the presence of only one Trp residue (28). Using known bimolecular quenching rates of Trp (27), our kinetic modeling using the rate model of Van Gilst and coworkers (28) indicates that the major contribution to fluorescence changes in h-apoMb comes from Trp-14 (A helix) and that its lifetime and quenching are most substantially affected by Met-131 (H helix), whose side chain is in van der Waals contact with the Trp ring (Fig. 3).

The kinetic modeling of Trp quenching includes rates for fluorescence, reversible bimolecular quenching by specific residues, and a background quenching rate (solvent and remainder of protein) (27, 28) using structural parameters for holo-Mb. In h-apoMb, Trp-7 is predicted to be heavily quenched (short lifetime) by neighboring Glu-6 and Gln-8 residues and shows no specific quenching outside the A helix; Trp-14 is weakly quenched by Ala-15 and Val-13 residues but

Table 1. Summary of observed lifetimes for the kinetic phase due to collapse

Glycerol (M)	T_i (°C)	T_f (°C)	η (cP)*	Nominal pH	$\tau_{collapse}$ (μs)	$\tau/(b\eta)^\dagger$
0	0 (2)	22 (3)	1.0	5.9	7 (5)	1.4 (9)
0	−7 (1)	10 (3)	1.3	5.2	5 (1)	0.8 (2)
0.1	−8 (1)	10 (3)	1.3	5.2	5 (3)	0.8 (4)
3	−12 (1)	5 (2)	3.5	5.2	17 (6)	1.0 (4)

Estimated errors are based on the reproducibility and variation found in several data sets taken under nominally identical conditions.

*Viscosity at final temperature T_f; see ref. 33.

†b = 4.89 is used to arbitrarily scale τ/η = 1 for the average of the four results.

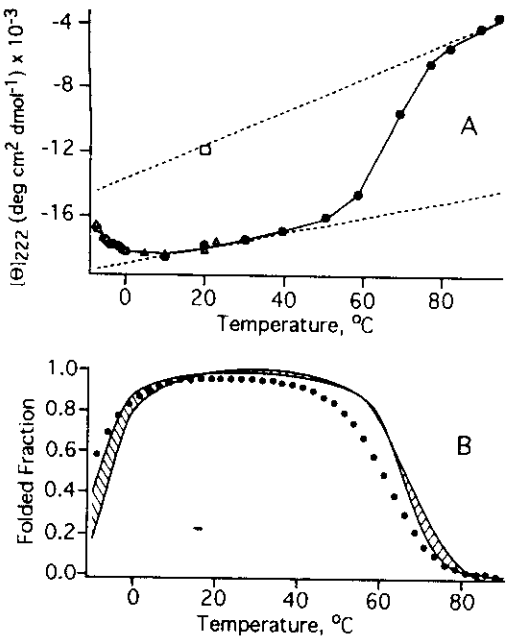


FIG. 2. (*A*) Mean residue ellipticity at 222 nm (disks) for h-apoMb in the −8–95°C range (pH 5.9). □, collected after the sample was held at −7.2°C for 1 h (fully reversible cold denaturation); ▽, the CD of mostly denatured protein (sample between 70–95°C for 1 h, then cooled to 20°C); solid line, a fit to a thermodynamic two-state model (21, 22), including fitted temperature-dependent baselines (dashed line). The unfolding thermodynamic parameters are ΔH = 195(17) kJ/mole, ΔS = 0.575(49) kJ/mole-deg, and ΔC_p = 5.1(5) kJ mole-deg. (*B*) The fraction of folded protein (the hatched zone represents the uncertainty due to thermodynamic parameters errors). The disks are the folded fraction in 3 M glycerol, which still shows substantial cold denaturation. (CD data for 0 and 3 M glycerol buffers at pH 5.2 are similar to those in *A* and not shown here; the uncertainty in the 3 M glycerol folded fraction is similar to that shown for 0 M, pH 5.9 conditions.) T-jumps in the −10 to 10°C range can produce 25–40% population changes.

shows strong specific quenching by Met-131 in the H helix (Met C_ε methyl to Trp C_{Z3} contact). The following control experiments (Fig. 4) support the modeling. Wild-type sperm whale apoMb has fluorescence properties that are very similar to those of h-apoMb (no differences in the critical residues described above). The Trp-7 → Phe-7 sperm whale mutant in Fig. 4 has a slightly larger slow component (locally quenched Trp-7 absent); guanidine·HCl-denatured h-apoMb has a large slow component with a smooth temperature dependence (no Trp-14–Met-131 quenching at any temperature, leaving only background quenching of the Trp); cold-denatured h-apoMb also has a large slow component, which rapidly decreases at higher temperature (Trp-14–Met-131 contact upon folding). Kinetic data (see below) on an Met-131-free mutant indicates that Met is indeed critical to Trp-14 quenching. A more detailed analysis with other residues is not warranted until uncomplexed apoMb structures become available, and it is not likely to affect our general conclusions.

FAST KINETIC RESULTS

Fig. 5 summarizes a fast folding experiment with h-apoMb, and conditions of several other experiments are listed in Table 1. The raw fluorescence data are analyzed by singular value decomposition (30) to eliminate noise components and by least-squares fitting. They show a clear progression from the cold-denatured value to compact-state, native values (as obtained on the same apparatus or a phase fluorimeter in a control after 0.5 ms or under steady-state conditions).

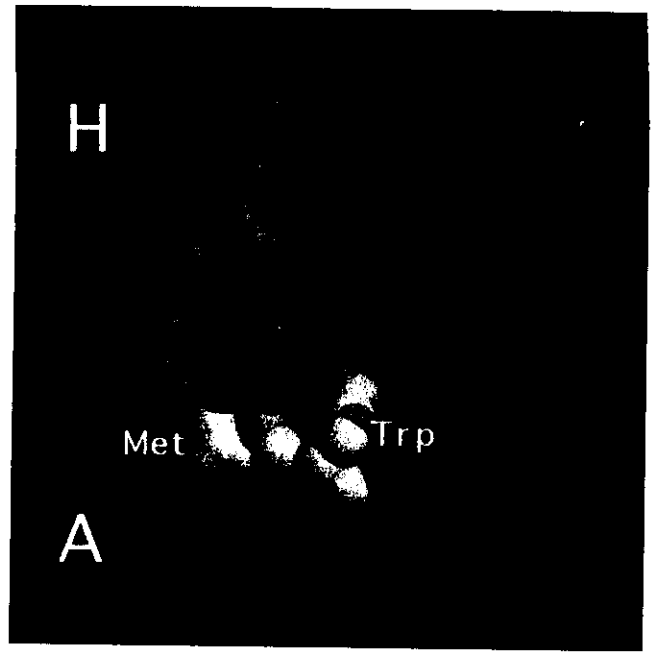


FIG. 3. A model of folded h-apoMb, based on ref. 23 (holoprotein), showing the Trp-14 and Met-131 residues involved in fluorescence quenching upon A–H helix contact. The color coding corresponds to theoretically derived folding units (29). (The folding units in this figure were derived from an energy landscape analysis by picking protein segments and then identifying segments with the largest ratio of stability gap to the spread of their molten globule energies.) G·H form such a unit (red) and the A helix is part of another unit (green), in which it shows the least structural fluctuations during folding simulations (15). This is highly suggestive of the postulated A·G·H folding intermediate (2) as well as the very fast collapse phase observed in our work.

We have employed a number of different methods to analyze the progression of individual fluorescence transients, including exponential fits with different floating parameters. Fig. 5*E* shows a typical result of fitting deconvoluted decays using a bi-exponential function. However, in this experiment, we are interested not in the details of individual fluorescence transients but in the progressive change of fluorescence as the protein folds. The most robust method of analysis thus bypasses the need for fluorescence lifetimes and instrument response functions entirely (Fig. 5*C*). An "initial" or "unfolded" fluorescence profile, f_1, and a "final" or "compact" profile, f_2, are generated from data before and long after the T-jump; a linear least-squares fit then reconstructs the data at intermediate times from the f_i by fitting their two amplitudes, A_1 and A_2. The fraction $\chi_2 = A_2/(A_1 + A_2)$ is plotted as a function of time delay. χ_2 is not very sensitive to residual amplitude fluctuations and mostly correlates with the shape (i.e., lifetime) of the fluorescence decays as they progress from cold-denatured to compact/native-like.

The kinetic behavior is typically as follows. Before the T-jump, $\chi_2 \approx 0$; within the next probe pulse (≤15 ns), χ_2 rises instantaneously to 0.8–0.9. This is followed by a rise of a few microseconds to $\chi_2 \approx 1$. The instantaneous change is nearly identical to a control experiment with free Trp or acetyl-Trp (Fig. 6). We assign it purely to the temperature dependence of the emission (local motion and population effects) rather than any ≈15-ns global folding motion. This is also verified by fluorescence lifetime measurements of h-apoMb at −8°C and 10°C in the presence and absence of denaturants (Fig. 4; *Steady-State Results*).

Fig. 6 illustrates another refolding transient obtained in 3 M glycerol solution at higher viscosity. The refolding amplitude is a larger fraction of the total amplitude in this case because

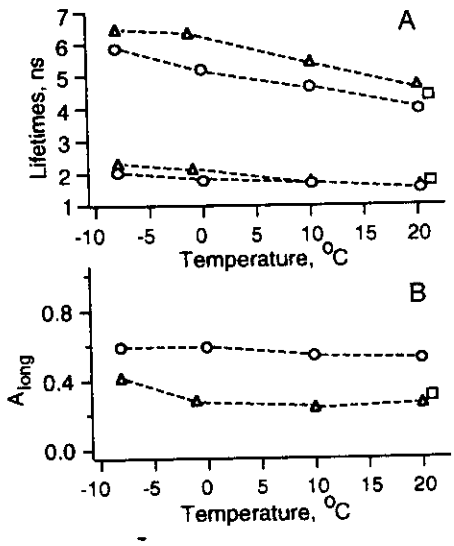


FIG. 4. Fluorescence properties of h-apoMb relevant to the fast folding experiment (pH 5.9); all lifetimes were fitted as bi-exponentials with a scatter component from phase fluorimetry data calibrated against *p*-terphenyl. (*A*) The short and long lifetime components in 0 M (△) and 3 M (○) guanidine·HCl. Lifetimes increase smoothly with decreasing temperature. The lifetimes of a sperm whale mutant with only Trp-14 present (□) are very similar to h-apoMb. (*B*) A_{long} (fractional amplitude of the long lifetime component) is nearly constant in 3 M guanidine·HCl, whereas it increases significantly upon unfolding in the absence of denaturant. The fractional amplitude of the long component for a sperm whale mutant without Trp-7 (□) is slightly larger than that for h-apoMb.

the initial temperature (−12°C vs. 0°C in Fig. 5) resulted in a larger initial fraction of unfolded protein (see CD in Fig. 2). These data will be discussed in detail below. Fig. 6*B* shows a control experiment performed with an aqueous Trp solution jumped by 22°C. Only the instantaneous phase due to the temperature dependence of the emission is evident in this test and similar tests, confirming that the microsecond phase is inherent to the protein-folding process. Furthermore, a kinetic trace (data not shown) obtained for an Met-131 → Ala-131 sperm whale apoMb mutant (purified sample kindly provided by R. Baldwin and M. Kay, Stanford University) refolding to the acid-globule state at pH 5.2 shows no microsecond phase at our signal-to-noise level, indicating that Met-131 is indeed the major quencher and that the A·G·H complex is likely not preformed in the cold-denatured state.

DISCUSSION

We assign the ≈7-μs kinetic phase in Fig. 5 to the time required for the unfolded ensemble (U) to form a compact (but still strongly hydrated) state. Using CD experiments with 5-ms time resolution, Jennings and Wright (2) have recently reported evidence for a molten globule (MG) state as a first step in apoMb folding. Their evidence is based on different (from native, N) urea titration curves and CD spectra of a <5-ms burst phase, which acquires 75% of the full apoMb optical rotation within the experimental dead time. While our fluorescence measurements cannot conclusively distinguish between a U → N or U → MG transition, it is very unlikely, based on diffusion and contact lifetimes for a 153-link polypeptide chain at ≈1 cP viscosity, that we are observing a very fast U → N phase. Rather, our 7-μs phase reflects the formation of a more compact state in which the Trp-14 and Met-131 residues have approached from a value closer to ≈40 Å (random coil value) to a ≈9 Å native C^{α}–C^{α} separation. This is in agreement with the expectation that the A·G·H complex of apoMb would be the earliest structure to form in the folding process, based on amide proton NMR protection data (2).

This collapse rate is slightly faster than expected from recent experiments by Eaton and coworkers (5, 6), which yielded diffusional collision times of residues in cytochrome *c* near 40 μs in the presence of denaturants. There are two points of view when comparing our results with theirs. The expected smaller roughness of the free-energy surface under their strongly denaturing conditions would point toward 40 μs being a lower limit on the collapse time; on the other hand, the shallower funnel under their denaturing conditions would point toward 40 μs being an upper limit. In any case, both numbers are close to the expected purely diffusional time constant, which should be of the same order as the collapse time for formation of a compact state. Our measured smaller time constant indicates that the A·H "collision" in apoMb occurs somewhat faster than expected from simple diffusion of a chain in a solvent, perhaps due to "bootstrapping" of the polypeptide chain by non-native hydrophobic contacts.

If the microsecond phase in Figs. 5 and 6 indeed leads from a largely unfolded coil to a more compact state with substantial native-like contacts in the A·G·H complex, one would expect a significant Kramers (i.e., diffusional) dependence of the rate on viscosity even at low viscosities (unless a convective motion "burrowing" through the solvent were involved). Ansari and coworkers (31) have measured the viscosity dependence of the structural rearrangement rate of folded Mb after photolytic detachment of CO. Their observed rearrangements are relatively small in scale compared with those during refolding, but they are globally distributed over the protein, particularly near the heme pocket. In the 0.7- to 3-cP viscosity range, their rate for conformational rearrangement in folded Mb is viscosity independent, whereas it decreases rapidly at higher viscosities. Their data are well-fitted by a modified Kramers expression that includes an internal protein friction term σ, such that the rate is proportional to $(\sigma + \eta)^{-1}$, where η is the solvent viscosity (31, 32).

We have measured the folding rate at 3.5-cP viscosity using 3 M glycerol/water solutions (33) of h-apoMb (Table 1). The resulting fit to a 17-μs kinetic phase is shown in Fig. 6. The refolding transient is considerably slower in this experiment than in 0 or 0.1 M glycerol experiments. This result agrees better with a strict Kramers-like decrease of the rate as a function of viscosity than with a constant rate, if solvent-induced changes in the activation energy can be neglected. [We have reason to believe that the 10°C and 22°C measurements in Table 1 indicate that the activation energy must be small to begin with: "stabilizing" solvents such as glycerol are known to be excluded from the immediate vicinity of the protein (34) and are likely to solvate an unfolded chain and hydrated transition state in a similar way, creating only a small change in activation barrier.]

The process in Figs. 5 and 6 is, therefore, due to solvent-exposed—most likely large-scale—motions of the protein backbone as it collapses to a compact state and not due to small-scale changes dominated by self-friction of a protein that has already largely excluded its aqueous environment. This is also supported by the CD data in Fig. 3; the data are well-fitted by a two-state model with the same baseline for the heat- and cold-denatured states. Therefore, our initial state should have largely coil-like characteristics with some residual helicity, unlike the native state or the structured A·G·H bundle inferred by Jennings and Wright for the molten globule (2).

Singular value decomposition data near $t = 0$ (average of several folding transients taken under nominally identical conditions, data not shown) with <20-ns dead time show a small kinetic phase leading to a longer lifetime immediately after $t = 0$, before the shortening due to Met-131 quenching. The lifetime of several hundred nanoseconds of this component may be due to incipient protection of Trp-14 from the

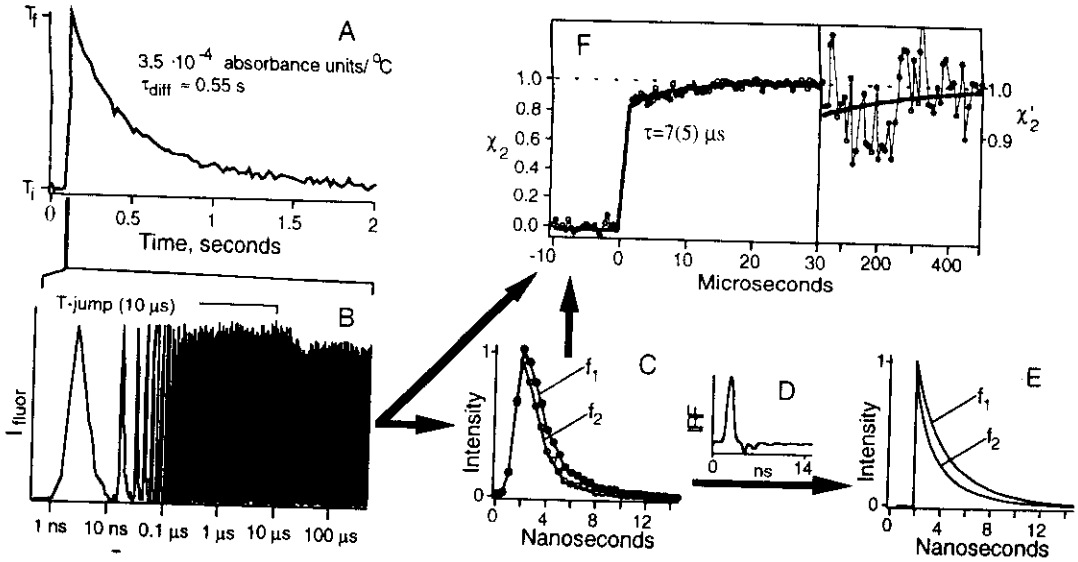


FIG. 5. Experimental data and kinetic fit. (*A*) Slow diffusional relaxation of cell temperature following the T-jump (see Table 1). On a logarithmic scale (*B*) covering six orders of magnitude in time, the baseline-subtracted fluorescence transients are shown from -10 μs to +490 μs (with respect to the T-jump as marked; pH 5.9, $3{\cdot}10^{-4}$ M protein). The T-jump is followed by an "instantaneous" phase and a 5- to 7-μs phase reaching the steady-state fluorescence. Individual transients (15-ns length) mark one point in the folding evolution. Depending on the time scale, data are averaged in 1–600 transient blocks to extract maximal signal-to-noise at long times. (*C*) To analyze the data, two fluorescence decay functions, f_1 and f_2, are generated by averaging the initial and final 8 μs of the interval. The data and f_i are singular value decomposed. The circles in *C* are constructed from two singular value decomposition components. Data can be fitted to a bi-exponential decay convolved with the instrument response function (*D*), which is precisely known from lifetime measurements using calibrated compounds (*p*-terphenyl) or 420-nm scattering as an instantaneous response. (*E*) The extracted decays, with $\tau_1 = 3.7$ ns, $A_1 = 60\%$, $\tau_1' = 1.2$ ns, $\tau_2 = 3.34$ ns, $A_2 = 43\%$, and $\tau_2' = 1.0$ ns. Generally, the lifetimes are not of interest, and a two-parameter linear least-squares fit of the data to the fixed functions f_1 and f_2 is used to determine their amplitudes, A_1 and A_2. (*F*) The resulting 7(5)-μs kinetic phase in $\chi_2 = A_2/(A_1 + A_2)$, which accounts for 15% of the signal, the rest being due to the instantaneous temperature effect. χ_2 is independent of overall fluorescence intensity and depends only on the shape (i.e., lifetime) of the signal as it progresses from unfolded toward the compact state signature. The right side of *F* shows the 30- to 490-μs time scale. Small phases of <0.5-μs and >100-μs duration may also be present but cannot be resolved with the present signal-to-noise ratio.

solvent by a newly formed A helix before its "collision" with the G·H complex, but better signal-to-noise and solvent quenching studies will be required to verify this.

Recent experiments by Nölting and coworkers (12) report a very small folding phase of ≈ 300 μs in T-jumped barstar. As shown in Fig. 5*F*, some of our runs show indications of a 100- to 500-μs phase that cannot be resolved with the present signal-to-noise ratio. If it is indeed present, it could be analogous to the early kinetics observed by Nölting and coworkers. One may speculate that this phase is due to smaller-scale rearrangements of the protein after initial collapse (with rearrangements of the hydration shell), as these would have a much less severe effect on the quenching of Trp fluorescence. Perhaps it is even a very fast phase leading to the native state (although data in ref. 12 indicate that it may still be solvent-exposed). In that case, our prediction is that this phase should show a minimal viscosity dependence over the 1- to 3-cP range considered here. Our faster phase would have been difficult to observe with their apparatus, if one is indeed present in barstar.

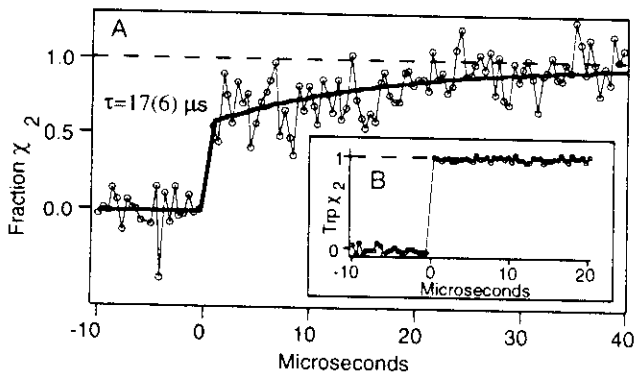


FIG. 6. (*A*) Kinetic transient for h-apoMb in 3 M buffered glycerol (pH 5.2). The 17(6)-μs phase of this sample [3.5 cP at 5°C (1 P = 0.1 Pa·s)] is significantly slower than observed in pure aqueous buffer (1.3 cP at 10°C). At low viscosity, this indicates that large-scale protein motions during folding dominate over internal protein friction effects. (*B*) Control experiment using Trp (pH 5.9, T-jump from 0 to 22°C). Only an instantaneous component is found in this and similar controls with acid-denatured h-apoMb and a Met-131 → Ala mutant of sperm whale apoMb.

In conclusion, our study with ≤20-ns dead time and time resolution shows that collapse of the unfolded state to a compact state involving at least the A- and H-helix backbone is complete in a few microseconds and involves significant motions of the backbone through the solvent. It remains to be seen from appropriate fluorescence/mutagenesis experiments whether other parts of partially folded apoMb (C·D·E·F) collapse more slowly or incompletely on this time scale. Future nanosecond time-resolution CD and infrared experiments will show how the timing of any postulated secondary structure nuclei compares to the collapse rate.

We thank P. G. Wolynes, Z. Luthey-Schulten, and S. Sligar for helpful discussions and critical comments, and E. Chien and S. Haire for help with sample preparation. This work was funded by the University of Illinois at Urbana–Champaign, Research Board, Grant 28526-AC7 from the grantors of the Petroleum Research Fund administered by the American Chemical Society, National Young Investigator Award of the National Science Foundation (Grant CHE-9457970), and the David and Lucile Packard Foundation. M.G. was also supported by a Cottrell Scholarship (Research Corporation) and a Camille and Henry Dreyfus New Faculty Award. Steady-state CD and fluorescence measurements were made at the University of Illinois Laboratory for Fluorescence Dynamics (supported by the National Institutes of Health and the University of Illinois at Urbana–Champaign).

1. Bychkova, V. Y. & Ptitsyn, O. B. (1993) *Biophysics (Engl. Transl.)* **38,** 51–58.
2. Jennings, P. A. & Wright, P. E. (1993) *Science* **262,** 892–896.
3. Schindler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995) *Nat. Struct. Biol.* **2,** 663–673.
4. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 9029–9033.
5. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J. & Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11860–11864.
6. Chan, C.-K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1996) *Biophys. J.* **70,** A177 (abstr.).
7. Williams, S., Cosgrove, T. P., Gilmanshin, R., Fang, K. S., Callendar, R. H., Woodruff, W. H. & Dyer, R. B. (1996) *Biochemistry* **36,** 691–698.
8. Phillips, C. M., Mizutani, Y., Hochstrasser, R. M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7292–7296.
9. Huang, G. S. & Oas, T. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6878–6882.
10. Pascher, T., Chesnick, J. P., Winkler, J. R. & Gray, H. B. (1996) *Science* **271,** 1558–1560.
11. Bernasconi, C. F. (1976) *Relaxation Kinetics* (Academic, New York). *contains references to the pioneering work of Eigen and coworkers, as well as early laser T-jump studies.*
12. Nölting, B., Golbik, R. & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci USA* **92,** 10668–10672.
13. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995) *Proteins Struct. Funct. Genet.* **21,** 167–195.
14. Dill, K. A., Bromberg, S., Yue, K. Z., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **4,** 561–602.
15. Goldstein, R. A., Luthey-Schulten, Z. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9029–9033.
16. Sali, A., Shakhnovich, E. & Karplus, M. (1994) *J. Mol. Biol.* **235,** 1614–1636.
17. Sharp, K. A., Nicholls, A., Friedman, R. & Honig, B. (1991) *Biochemistry* **30,** 9686–9697.
18. Teale, F. (1959) *Biochim. Biophys. Acta* **35,** 543.
19. Wetlaufer, D. B. (1962) *Adv. Protein Chem.* **17,** 303–390.
20. Egeberg, K. D., Springer, B. A., Martinis, S. A., Sligar, S. G., Morikis, D. & Champion, P. M. (1990) *Biochemistry* **29,** 9783–9791.
21. Privalov, P. L., Griko, Y. V., Venyaminov, S. Y., Kutyshenko, V. P. (1986) *J. Mol. Biol.* **190,** 487–498.
22. Nishii, I., Kataoka, M. & Goto, Y. (1995) *J. Mol. Biol.* **250,** 223–238.
23. Evans, S. V. & Brayer, G. D. (1990) *J. Mol. Biol.* **213,** 885–897.
24. Loh, S. N., Kay, M. S. & Baldwin, R. L. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5446–5450.
25. Bismuto, E., Gratton, E. & Irace, G. (1988) *Biochemistry* **27,** 2132–2136.
26. Nash, D., Lee, B. & Jonas, J. (1996) *Biochim. Biophys. Acta*, in press.
27. Steiner, R. F. & Kirby, E. P. (1969) *J. Phys. Chem.* **73,** 4130–4135.
28. Van Gilst, M., Tang, C., Roth, A. & Hudson, B. (1994) *J. Fluorescence* **4,** 203–207.
29. Panchenko, A. R., Luthey-Schulten, Z. & Wolynes, P. G. (1996) *Proc. Natl. Acad. Science USA* **93,** 2008–2013.
30. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992) *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, U.K.), 2nd Ed., pp. 51–63.
31. Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1992) *Science* **256,** 1796–1798.
32. Hagen, S. J., Hofrichter, J. & Eaton, W. A. (1996) *Biophys. J.* **90,** A178 (abstr.).
33. Dean, J. A., ed. (1985) *Lange's Handbook of Chemistry* (McGraw-Hill, New York), 13th Ed.
34. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22,** 67–97.