Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., & Russell, D. W. (1984) *Cell 39*, 27–38.
Yuan, D., & Tucker, P. (1984) *J. Exp. Med. 160*, 564–583.
Zaphiropoulos, P. G., Mode, A., Ström A., Möller, C., Fernandez, C., & Gustafsson, J.-Ä. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4214–4217.
Zhao, H.-F., Simard, J., Labrie, C., Breton, N., Rhêaume, E., Luu-The, V., & Labrie, F. (1989) *FEBS Lett. 259*, 153–157.
Zhao, H.-F., Labrie, C., Simard, J., de Launoit, Y., Trudel, C., Martel, C., Rhêaume, E., Dupont, E., Luu-The, V., Pelletier, G., & Labrie, F. (1991) *J. Biol. Chem. 266*, 583–593.

Accelerated Publications

Detection and Characterization of an Early Folding Intermediate of T4 Lysozyme Using Pulsed Hydrogen Exchange and Two-Dimensional NMR†

Jirong Lu and Frederick W. Dahlquist*

Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oregon 97403-1229

Received January 10, 1992; Revised Manuscript Received March 27, 1992

ABSTRACT: Two-dimensional ^{1}H–^{15}N NMR techniques combined with pulsed hydrogen–deuterium exchange have been used to characterize the folding pathway of T4 lysozyme. In the unfolded state, there is little differential protection of the various amides from hydrogen exchange. In the native folded structure, 84 amides of the 164 residues are sufficiently spectrally resolved and protected from solvent exchange to serve as probes of the folding pathway. These probes are located in both the N-terminal and C-terminal domains of the native folded structure of the protein. The studies described here show that at least one intermediate is formed early during refolding at low denaturant concentrations. This intermediate (or intermediates) forms very rapidly (within the 10-ms temporal resolution of our mixing device) under the conditions used and is completed at least 10 times faster than the overall folding event. The intermediate(s) protect(s) from exchange a subset of amides in the N-terminal and C-terminal regions of the protein. In the final folded states these protected regions correspond to two α-helices and a β-sheet region. These amides are protected from exchange by factors between 20 and 200 as compared to the fully unfolded protein. Protection of this magnitude is consistent with the formation of somewhat exposed secondary structure in these regions and could represent a "molten globule"-like or a "framework"-like structure for the intermediate(s) in which specific parts of the sequence form isolated secondary structures that are not stabilized by extensive tertiary interactions.

The lysozyme produced by T4 bacteriophage is a monomeric protein with no disulfide bonds. Figure 1 shows the amino acid sequence of T4 lysozyme with secondary structure segments denoted as helix A, B, C, D, E, F, G, and H and β-sheets determined by both crystallographic and solution methods (Matthews & Remington, 1974; McIntosh et al., 1990). The N-terminal domain (13–71) has more β-sheet-like structure, and the C-terminal domain (1–12, 72–164) is largely α-helical. The two domains are linked in part by a long helical segment (helix C).

The folding and stability of this 164-residue protein have been the subject of extensive study using X-ray crystallographic and solution methods (Alber et al., 1987; Becktel & Baase, 1987; Hudson et al., 1987; Matthews, 1987; Weaver & Matthews, 1987). A large collection of point mutants has been generated to investigate the structural and thermodynamic consequences of modification of particular residues in the protein [see Bell et al. (1990)]. Under most conditions, the equilibrium between the folded and unfolded states of most of the lysozyme variants is two-state, showing essentially no significant equilibrium population of any intermediate species. Similarly, the recent kinetic studies of Schellman and coworkers show no evidence for the accumulation of kinetic intermediates (Chen et al., 1991) in the folding pathway in the presence of high concentrations of denaturants. These elegant studies were designed to investigate the folding pathway in a regime where the relaxation times are accessible using manual mixing techniques. The denaturant concentrations needed to slow the relaxation rates to this extent reveal very interesting interactions of the protein with solvent but mask events that occur under conditions where the native form is most stable.

We were particularly interested in the possibility that we might detect and begin a structural analysis of early kinetic intermediates in the folding pathway. In this event, the observed intermediate(s) would reflect a population of protein molecules in which there are no obvious slow kinetic barriers such as hindered rotation of a peptide bond separating members of the majority population. The kinetic properties of such a population are likely to provide interesting insights into the early events in folding. Structural information about the

†This work was supported by grants from the National Science Foundation (DMB8905322 and DIR8815047), the NIH (1S10RR04926), and the Markey Charitable Trust.
*Author to whom correspondence should be addressed.

0006-2960/92/0431-4749$03.00/0

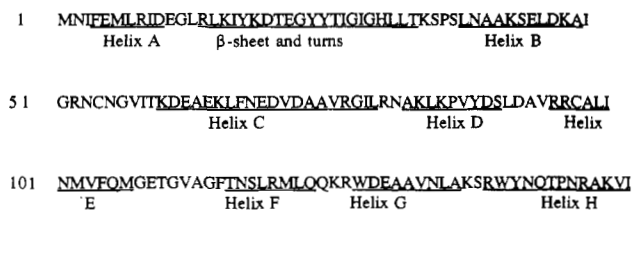

FIGURE 1: Diagram of the amino acid sequence of T4 lysozyme with secondary structure segments indicated.

properties of such intermediates is essential for understanding the mechanism of protein folding.

Recently, two-dimensional NMR techniques combined with quench-flow measurements of hydrogen–deuterium exchange of amide protons have been used to detect and characterize kinetic intermediates observed during protein folding. This method has been applied to RNase A, cytochrome *c*, and barnase (Udgaonkar & Baldwin, 1988; Roder et al., 1988; Bycroft et al., 1990). This approach can also provide some structural information about intermediates on the folding pathway since measurement of hydrogen–deuterium exchange rates of amide protons is a sensitive probe for the presence of secondary structure in a fully or partially folded protein [see Baldwin and Roder (1991)]. The formation of hydrogen bonds such as those seen in α-helices or β-sheets dramatically reduces hydrogen–deuterium exchange rates of amide protons. In a typical pulsed hydrogen-exchange experiment, refolding of a deuterated protein is initiated in D_2O, and a snapshot of the protein is taken at different times after refolding by subjecting the sample to H_2O at high pH to label those amides that remain exposed. Two-dimensional NMR is then used to determine those amide protons that are protected at various times during the folding process.

We recently assigned the backbone ^1H and ^{15}N NMR resonances of T4 lysozyme (McIntosh et al., 1990) using a combination of selective and nonselective biosynthetic incorporation of isotopes into the protein. The availability of the backbone ^1H–^{15}N assignments allows us to use the pulsed hydrogen-exchange experiment described above to investigate the folding pathway of this protein.

Materials and Methods

T4 Lysozyme Preparation. The two cysteine residues of T4 lysozyme (Cys54/Cys97) were replaced by threonine and alanine, respectively, to increase the reversibility of folding and unfolding at high pH values. Both the stability and structure of this protein are similar to those of wild-type T4 lysozyme. This variant is referred to as WT* (Anderson et al., 1990; Matsumura & Matthews, 1989). The genes encoding WT* and the T4 lysozyme mutants were cloned into an inducible plasmid system for high-level expression in a variety of *Escherichia coli* strains. T4 lysozyme with uniform ^{15}N enrichment was produced using a prototrophic *E. coli* strain grown in M9T media containing 1.0 g of $(^{15}NH_4)_2SO_4$ (Isotec)/L as the sole nitrogen source (Muchmore et al., 1989).

Urea-d_4 was obtained from MSD Isotopes.

Stopped-Flow Kinetics. Approximately 1 mg/mL protein was dissolved in 3.0 M urea, 25 mM potassium chloride, and 20 mM potassium phosphate at pH 2.0, 23 °C, for 2 h. Under these conditions the protein is completely unfolded. Refolding was initiated by diluting this solution 1:1 with a solution of 200 mM potassium phosphate, pH 7.0 at 23 °C. Refolding was monitored by the decrease of tryptophan fluorescence, measured in a Dionex D-137 stopped-flow fluorometer (deadtime about 3 ms) with excitation at 280 nm. Emission was measured with an O-54 cutoff filter. Data were recorded on a Nicolet Model 206 digital oscilloscope and plotted with an HP 7224 plotter. The rate constants and amplitudes of the fast and slow phases were obtained by fitting the time-dependent fluorescence change to a double exponential of the form $V(t) = Ae^{-k_1t} + Be^{-k_2t} + V_\infty$, where A and B are the amplitudes of the fast phase and slow phase, respectively, k_1 and k_2 are apparent rate constants for the fast phase and slow phase of refolding, respectively, and V_∞ is the final fluorescence. The Plotdata (Triumf, 4004 Wesbrook Mall, Vancouver, BC, Canada V6T2A3) package was used to analyze the data.

Pulsed Hydrogen-Exchange Experiments. Experiments were done at 20 °C using a rapid (millisecond) mixing quench-flow machine from Update Instruments Inc. Purified proteins were exchanged into D_2O at pH 1.7 through a Sephadex G-25 spin column. This protein was then unfolded in a buffer consisting of 3.0 M urea, 25 mM sodium chloride, and 20 mM H_3PO_4 in D_2O, final pH 2.7. This unfolded solution (15 mg/mL) was allowed to stand overnight at 20 °C to complete amide hydrogen exchange. The unfolded T4 lysozyme solution was diluted 1:1 into a refolding buffer consisting of 100 mM sodium acetate in H_2O, pH 6.0, to initiate refolding. At different times after the initiation of refolding, the protein solution was diluted 1:1 into the exchange probe buffer consisting of 100 mM glycine, 50 mM sodium phosphate, and 1.5 M urea at pH 9.5. After 41 ms, amide hydrogen exchange in 16% deuterated buffer at pH 9.5 was quenched by diluting the sample 1:2.5 into the quench buffer solution consisting of 100 mM potassium chloride and 50 mM H_3PO_4 on ice such that the final pH was 3.2. The refolding reaction was allowed to go to completion under these conditions. The samples were then concentrated with Centriprep concentrators to a volume of about 0.5 mL and prepared for NMR analysis.

Hydrogen–Deuterium Exchange in the Unfolded State. Protein at 15 mg/mL was unfolded in 8 M urea, 100 mM potassium chloride, and 50 mM potassium phosphate, pH 7.5 at 20 °C, in D_2O overnight. The rapid quench-flow device was used for the following treatments of the sample. Exchange was initiated by diluting the solution 1:2 with the same buffer solution in H_2O. After various times of incubation in 67% H_2O, exchange was quenched by diluting the sample 10-fold by injection into a vigorously stirred quench buffer (100 mM potassium chloride, 50 mM potassium phosphate, pH 3.0) on ice.

NMR Sample Preparation. A reference NMR sample was prepared by heating a 20 mg/mL solution of protein in 16% D_2O and 25 mM potassium chloride, about pH 2.0 at 50–60 °C, for 2 h to unfold the protein and exchange about 16% of the NH protons so that all the amide sites were about 84% protonated. The sample was then cooled to refold the protein. This sample was used as an intensity standard for the volume integrals used to quantitate proton occupancy.

Samples used in pulsed hydrogen-exchange experiments and hydrogen–deuterium exchange in the unfolded state were concentrated with Centriprep concentrators to 0.5 mL after the quench step from a total volume of about 15 mL.

All the samples were exchanged into D_2O buffer containing 100 mM potassium chloride and 50 mM potassium phosphate at pH* 3.0 by passing approximately 0.5 mL of the protein solution through a 2.5-mL Quik-sep spin column containing Sephadex G-25 preequilibrated in the deuterated buffer.

NMR Experiments. The HSMQC spectra were recorded at 500 MHz on a General Electric GN-500 spectrometer as

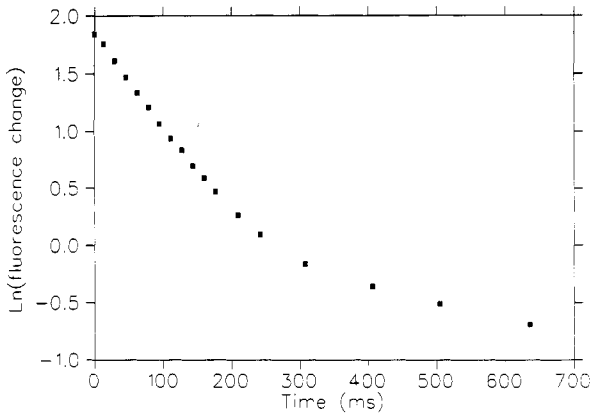

FIGURE 2: Semilogarithmic plot of the time-dependent fluorescence change during refolding of WT* T4 lysozyme in 110 mM potassium phosphate, 12.5 mM potassium chloride, and 1.5 M urea at 23 °C, pH 6.8.

previously described (McIntosh et al., 1990). The spectrum was recorded with a ^1H sweep width of 6666.7 Hz. Each free induction decay consisted of 1024 complex data points, zero filled to 2048, and 32 scans were collected. The ^{15}N sweep width was 3225 Hz. A total of 512 complex points in the t_1 dimension were collected. The total data acquisition time was 3.8 h. The FTNMR program of Dennis Hare was used for data processing on a Microvax 3 computer. Before Fourier transformation, the time–domain data were multiplied by a Gaussian window function, and t_1 domain data were extended to 1024 points with zero filling.

Volume integrals were obtained by manually choosing the peak center and base for each of the 84 amides using the reference sample spectrum. The volumes of the amide peaks were determined using the base area chosen for the reference sample spectrum. Volumes were than obtained for each time point using FTNMR to generate an array of peak volumes and times. The volumes were scaled using the resolved methyl resonance of valine 94. The scaled volumes as a function of time were then fit to a single exponential of the form $V(t) = V_0e^{-kt} + V_\infty$, where $V_0 + V_\infty$ is the initial volume, V_∞ is the final volume of the peak, and k is the apparent rate constant for the process.

Results

Overall Refolding Kinetics Monitored by Fluorescence Change. Refolding of T4 lysozyme was first monitored by the change of the fluorescence of its three tryptophan residues (Trp126, Trp138, and Trp158) to find the appropriate time scale for pulsed hydrogen-exchange experiments. The protein was denatured in 3 M urea, 25 mM potassium chloride, and 20 mM potassium phosphate at pH 2.0, and refolding was initiated by a combined pH and urea concentration jump to a final concentration of 1.5 M urea at pH 6.8, 23 °C. As shown in Figure 2, the fluorescence-detected refolding time course has two kinetic phases. The fast phase has a relaxation time of 120 ms, accounting for 90% of the total amplitude. The slow phase has a relaxation time of 650 ms, accounting for 10% of the total amplitude. The low amplitude of the slower phase suggests a potential role for proline isomerization in this phase. To test this idea, the three proline residues of WT* were replaced by alanine by site-specific mutation. The resulting triple mutant protein, P37A/P86A/P143A, has thermodynamic stability similar to that of the original and is fully active. The fluorescence-detected refolding time course of this protein lacking proline residues showed only the faster phase relaxation (data not shown). We conclude that one or more proline residues are responsible for the slow phase of refolding under these conditions. The role of proline isomerization in the slow step will be discussed elsewhere in detail (Lu and Dahlquist, in preparation).

Assignment of the HSMQC ^1H–^{15}N Spectrum of WT.* The NMR spectrum of WT* revealed that there were several resonances that change because of the cysteine substitutions. This required that the spectrum of the substituted protein be carefully reassigned. Peaks with little or no shift in the HSMQC ^1H–^{15}N spectrum of WT* were tentatively assigned on the basis of the assignment of wild-type T4 lysozyme (McIntosh et al., 1990). Assignments of WT* were completed and confirmed by edited NOESY, COSY experiments of uniformly labeled protein and with protein selectively labeled by growth with [^{15}N]glutamate and [^{15}N]aspartate. This results in labeled glutamate, aspartate, glutamine, asparagine, and arginine residues with lower efficiency labeling of alanine and threonine residues (McIntosh et al., 1990). The resulting assignment of the HSMQC ^1H–^{15}N of WT* is shown in Figure 3.

Hydrogen-Exchange Properties of the Unfolded State. Using the assignments shown in Figure 3, we could characterize the hydrogen-exchange properties of the unfolded state. The observed time courses for replacement of deuterium by hydrogen at the 84 individual amide sites used as probes were all adequately described by a single first-order decay. The observed first-order rate constants for exchange of each residue were then corrected to account for the pH of the sample, the presence of urea (Loftus et al., 1986), and the effects of adjacent residues to generate a calculated protection factor (Molday et al., 1972) for each residue. The protection factor should be near unity if a particular amide has the exchange properties of a fully solvent exposed amide. Values greater than unity reflect an amide whose rate of solvent exchange has been slowed by a factor equal to the protection factor as the result of the formation of structure(s) that inhibit exchange [see Baldwin and Roder (1991)].

The results are summarized in Figure 4. There is very little difference in protection factor as a function of position, and the value is within a factor of 5 of unity. The only significant apparent protection is seen at the various isoleucine and valine residues. Similar slower exchange of valine and isoleucine residues in the thermally unfolded state of RNase has been observed by Robertson and Baldwin (1991). Using model peptides, they have shown that fully exposed valine residues have intrinsically slower hydrogen exchange than other residues and the apparent inhibition of exchange at these positions does not reflect the formation of secondary structure. There are several residues that have protection factors less than unity. This was also observed in the unfolded state of RNase. Robertson and Baldwin suggested that this effect may be the result of electrostatic interactions in the vicinity of the particular amides. Our data are consistent with this view. The data shown in Figure 4 suggest that the unfolded state produced under these conditions has very little structure and that essentially all the amides are completely exposed to solvent.

Folding Kinetic Patterns of T4 Lysozyme Measured by the Pulsed Hydrogen-Exchange Experiment. The time scale for the pulsed hydrogen-exchange experiments was chosen to correspond to that of the rapid (t = 120 ms) phase of the refolding seen in Figure 2. Pulsed hydrogen-exchange experiments with WT* were done at seven different refolding times between 8 and 400 ms. Unfolded protein in which the amide protons had been exchanged for deuterium was allowed to refold in 50% D_2O at pH 5.1 for a variable length of time

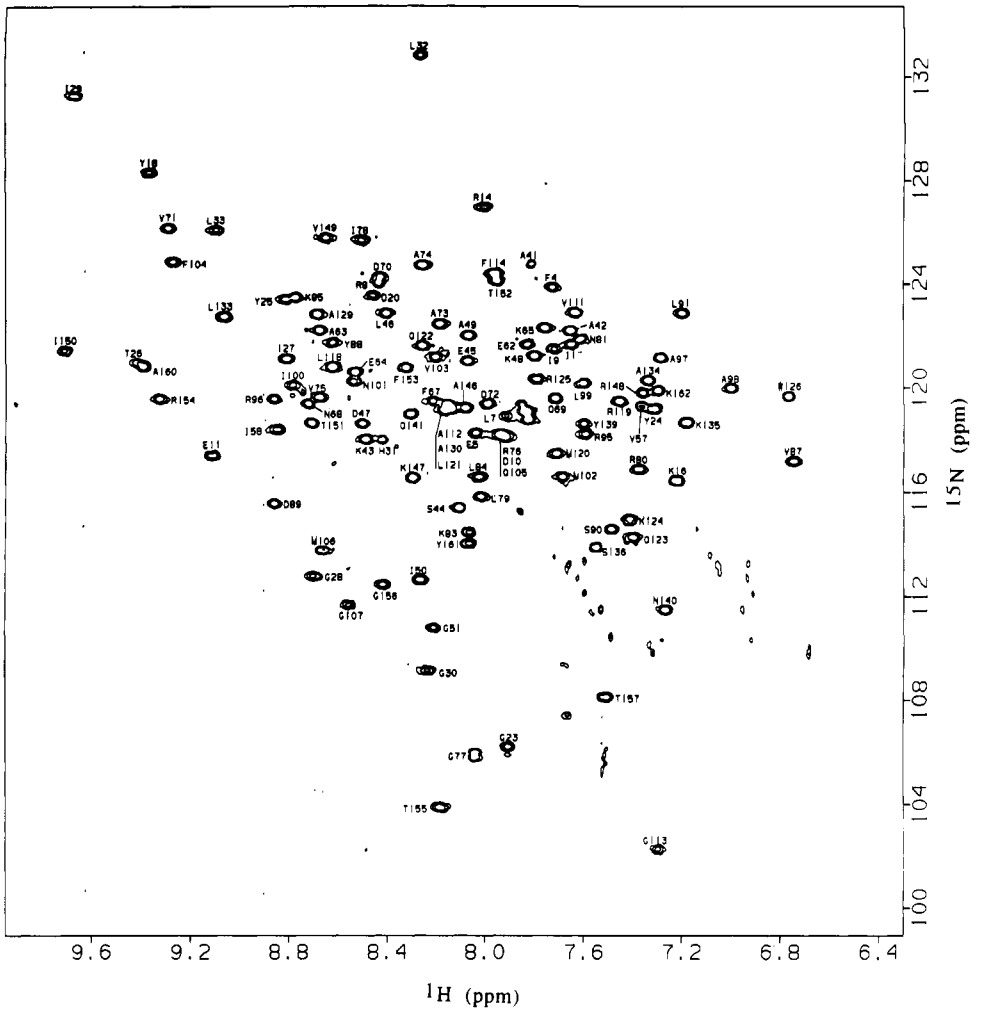

FIGURE 3: Reference HSMQC spectrum of uniformly ^{15}N-enriched WT* T4 lysozyme in D_2O, pH 3.0 at 21.5 °C. The amide protons depicted here are stable to exchange with solvent deuterons. The assignments of backbone amide protons are indicated.

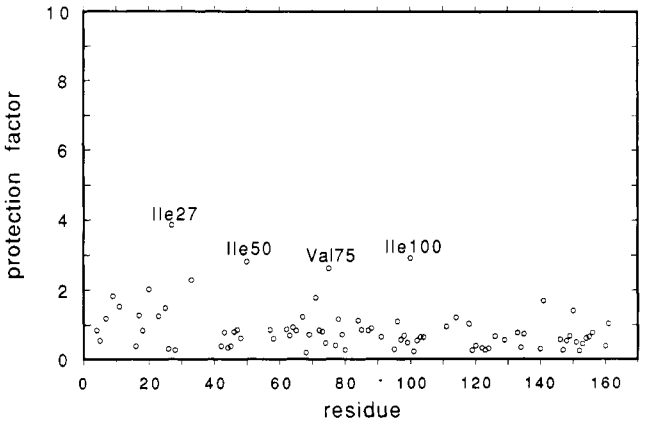

FIGURE 4: Plot of the protection factors as a function of residue number for the unfolded state of T4 lysozyme produced by treatment with 8 M urea, 100 mM potassium chloride, and 50 mM potassium phosphate, pH 7.5 at 20 °C.

(between 8 and 400 ms) and then exposed to a 41-ms proton-labeling pulse at pH 9.0. Exchange at pH 5.1 is negligible in the period of refolding. The exchange time constant for a free amide at pH 9.0 is about 1 ms (Molday et al., 1972). Amide protons in the unstructured parts of protein become fully protonated by the pH 9.0 pulse-labeling period, and the proton label is excluded from sites where exchange is retarded more than 50-fold by prior formation of structure. This labeling pattern established at various times along the folding pathway is transferred to the native protein and is fixed by quenching hydrogen–deuterium exchange at pH 3.2. At this pH, amide proton exchange is halted but refolding can proceed. After refolding was completed, the protein samples were concentrated and exchanged to the same buffer as that used for recording the reference spectrum, and the HSMQC ^{1}H–^{15}N spectra were recorded. Figure 5 shows a region of the HSMQC spectra of a reference sample and three representative samples labeled at different refolding times. The upper level of cross-peak intensities was determined from the reference sample that was prepared by unfolding the fully protonated protein in 16% D_2O at pH 2.0 for 2 h. A control experiment using the fully folded protein was done to test the resistance of individual amide protons in native T4 lysozyme to hydrogen-exchange labeling. In this experiment the fully folded protein was subjected to the same pH 9.0 proton-labeling pulse, quench, and concentration followed by observation of the HSMQC spectrum. This experiment confirmed that the proton label was completely excluded from all amide sites that are used as probes.

As shown in the spectra of Figure 5, there is a substantial difference in the time course of proton occupancy for different sites. This is shown more clearly in Figure 6 as a plot of the relative proton occupancy as a function of refolding time for the amides of Ile27, Asn68, Ile100, and Ala129. These four residues represent the range of behaviors seen for all probe residues. At the first time point, 8 ms, these residues have become protected against exchange to different degrees. Like most of the 84 probes, Asn68 and Ala129 become slowly protected against exchange with an apparent time constant

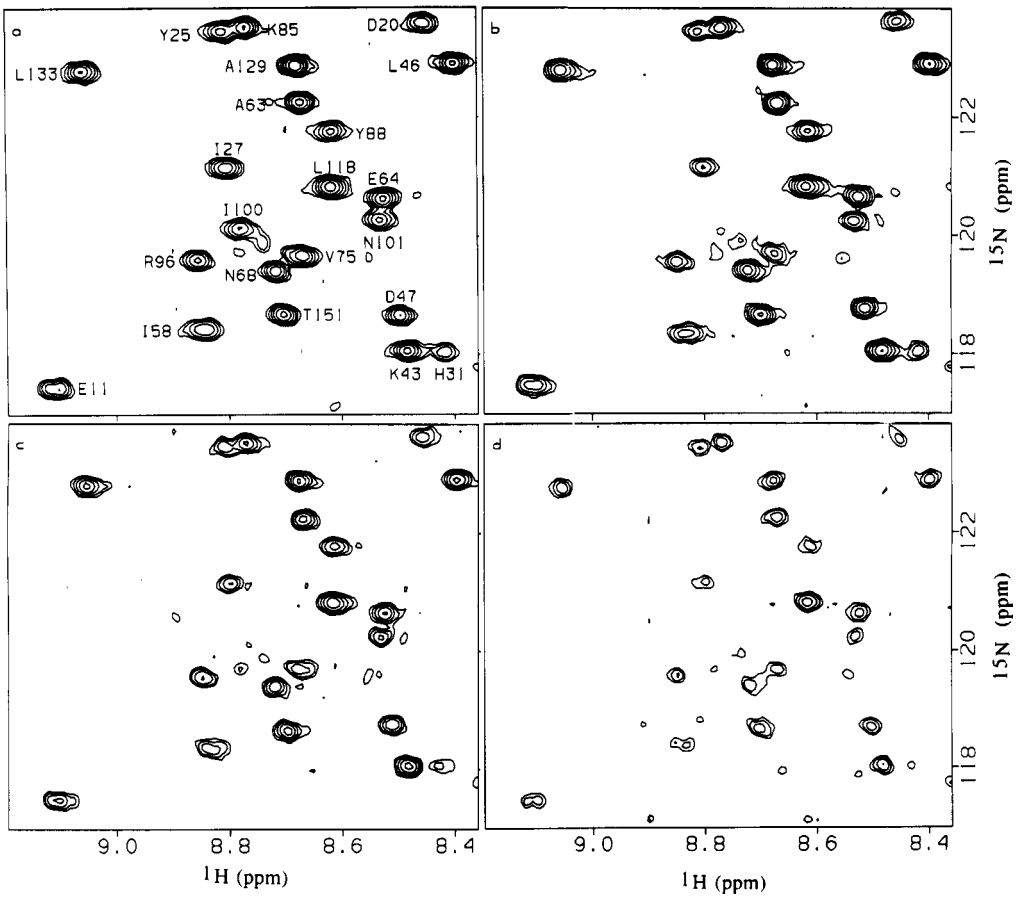

FIGURE 5: Two-dimensional NMR spectrum of representative WT* T4 lysozyme samples prepared by pulsed hydrogen-exchange methods: (a) reference sample; (b) 8-ms sample; (c) 100-ms sample; (d) 250-ms sample.

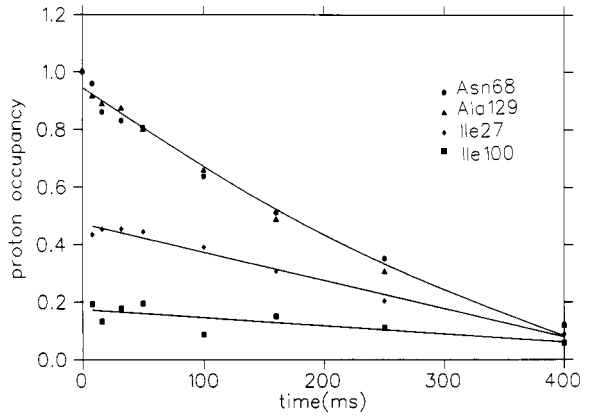

FIGURE 6: Time courses for the protection of backbone amides against exchange. Smooth lines were used to connect data for each residue.

of about 170 ms. This apparent time constant represents that of overall refolding. It is longer than the overall refolding time constant measured by the stopped-flow fluorometer due to slightly different conditions. At 8 ms these residues are almost completely accessible to solvent and are strongly labeled by protons in the 41-ms pulse of pH 9.0 used. These residues reflect regions that have not formed a structure capable of protecting those amides from exchange.

The opposite behavior is represented by Ile100. Here the amides of this residue are completely protected against exchange by the high pH pulse even at the earliest time. This result reflects the formation of some early structure that is capable of essentially complete protection of residue 100 against exchange at this pH. Similar behavior is shown by Val103.

The behavior shown by Ile27 is intermediate. At early times, this residue is partially protected from exchange during the pulse. Complete protection occurs with a time constant of about 170 ms. This intermediate behavior is observed for roughly one-third of the probe residues with varying degrees of protection at early times. There are two potential explanations for the intermediate behavior shown by this class of residues. In one, all the molecules follow essentially the same folding pathway, and the partial proton occupancy observed reflects the solvent-exchange properties of the residue in the intermediate. For Ile27, the high pH pulse would then be sufficient to cause about 50% exchange during the 41-ms duration of the pulse. In the second explanation, the folding pathway would be heterogeneous. Here the 50% proton occupancy observed for Ile27 at early times would be explained by a mixed pathway. About 50% of the molecules would form an intermediate or intermediates in which Ile27 would be strongly protected, and the remaining molecules would not form this intermediate or intermediates and would not be protected from solvent exchange.

To resolve this question, the pH of the pulse was varied. Proton occupancy at each site is determined by both the stability of the intermediate(s) at that site and pulse intensity (Udgaonkar & Baldwin, 1990; Elöve & Roder, 1991). Figure 7 shows the relative proton occupancy observed for several representative residues (Lys16, Tyr88, Ala97, Ile100, Trp126) when a 41-ms variable pH pulse was applied after 32 ms of refolding. At pH 9.55, most residues are about 85% labeled with protons after 32 ms of refolding except for a few residues in the N-terminal β-sheet (Lys16, Ile17, Tyr18, Ile27) and helix E (Arg95, Leu99, Ile100, Asn101, Met102, Val103, Phe104) showing some degree of protection. At pH 9.0, residues in helix A and the β-sheet in the N-terminal lobe and helix E in the C-terminal lobe show more protection, and a few residues in helix C that connects the N-terminal lobe with the C-terminal lobe show some protection.

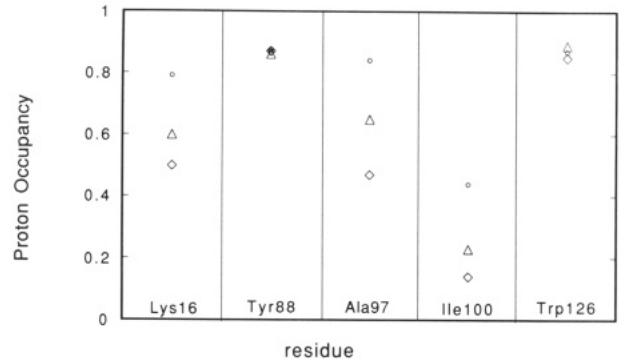

FIGURE 7: Degree of protection for five representative residues at three different pH values. The protein samples were refolded for 32 ms before being exposed to a 41-ms pulse at three different pH values. Symbols: (○) pH 9.55 measurement; (△) pH 9.00 measurement; (◇) pH 8.70 measurement.

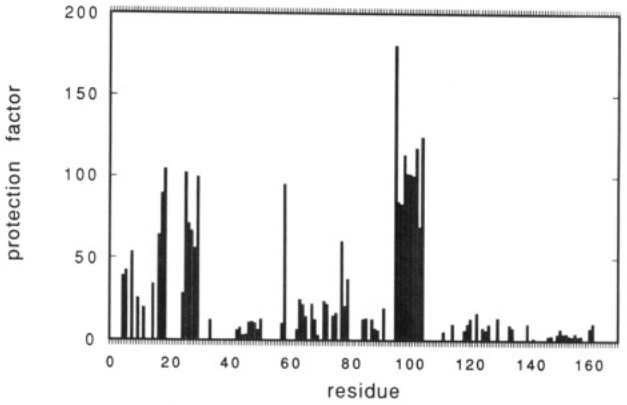

FIGURE 8: Plot of the protection factor of the intermediate(s) as a function of residue number.

The extent of protection varies for different residues and varies with the pH of the pulse. For those residues that show protection from solvent exchange at early times, the proton occupancy observed increases as the pH of the pulse is increased. This trend (as seen for Ala97) continues until the proton occupancy seen at early times approaches that of an unprotected residue such as Tyr88. These observations strongly suggest that the partial proton occupancies seen at early times for residues such as Ile27 are not the result of a bifurcation of the folding path. Rather, it appears that most if not all the molecules form an intermediate that is characterized by moderate protection from solvent exchange.

The base-catalyzed rate constant for solvent exchange, k_{OH}, can be calculated for those residues that show a proton occupancy between 0.2 and 0.8 following at least one high pH pulse treatment. If no exchange were to occur during any high pH pulse, the observed proton occupancy would be 0.2 due to the proton content of the original deuterated solvent used to produce the unfolded protein. If the only event that leads to protection is the 170-ms process, the observed proton occupancy would be 0.83 after 32 ms of folding has occurred. The proton occupancy, p, observed following 32 ms of refolding and a 41-ms pulse of water at a given concentration of hydroxide ion, [OH], is given by

$$p = 0.20 + 0.63\{1 - \exp(-0.041 k_{OH}[OH])\}$$

The value of k_{OH} was calculated for the residues that show a proton occupancy between 0.2 and 0.8 following at least one of the high pH pulses. This information is converted into a more useful form in Figure 8 as a plot of relative protection factor versus position. Here the relative protection factor was calculated as the ratio of the based-catalyzed rate constants for the residues of the intermediate(s) to the base-catalyzed rate constants of the unfolded state calculated from the data shown in Figure 4 with correction of the effect of urea on exchange rate. For residues such as Y88 and W126 (Figure 7) that did not show any protection at pH 9.55, 9.0, and 8.7, protection factors were calculated from additional measurements done at pH 8.02, 7.75, and 7.53. The calculated values of the base-catalyzed rate constants for exchange of the intermediate(s) are complicated by the possibility that both refolding and exchange have occurred during 41-ms pulse. We treated the proton occupancies as if no significant refolding had occurred during the pulse. This could lead to somewhat larger values of the protection factor than actually apply.

These data clearly show that three regions are protected from amide exchange within 8 ms of folding. The calculated relative protection factors are at least 20 and vary to a maximum of 180 for the residues in three regions of the sequence. This is shown schematically in Figure 9. The 84 amides that were used as probes are indicated by filled circles. Residues with a protection factor of 20 or higher are depicted in Figure 9 with larger filled circles. They correspond to β-sheets in the N-terminal domain, helix A and helix E in the C-terminal domain, and a few residues in helix C.

Discussion

We have employed pulsed hydrogen-exchange techniques to examine events along the folding pathway of T4 lysozyme. The unfolded state in 8.0 M urea, pH 7.5, 100 mM potassium chloride, and 50 mM potassium phosphate apparently lacks a measurable amount of stable secondary structure. A similar pattern is seen when the experiment is carried out in 3.0 M urea and pH 2.7 (data not shown), corresponding to the initial

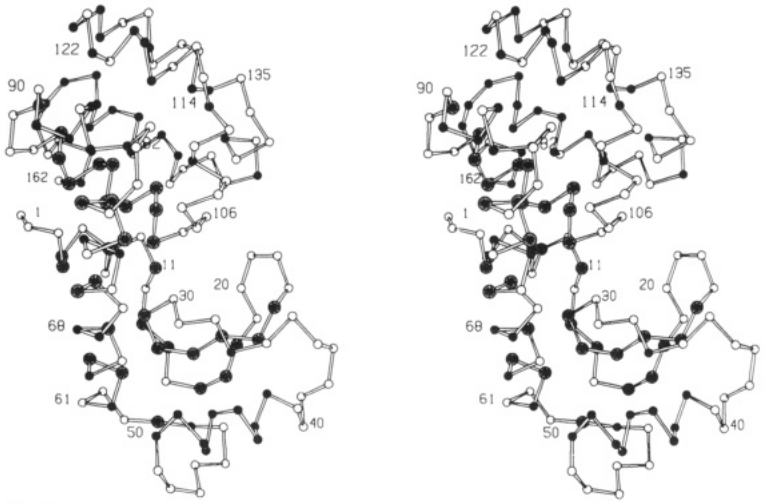

FIGURE 9: Tertiary structure of T4 lysozyme. The probe residues are shown with filled circles, and those residues with protection factor 20 or higher are labeled with larger filled circles. Open circles are α-carbons for each residue.

unfolded conditions used in the pulsed hydrogen-exchange kinetic experiments. The results strongly suggest that there is no stable secondary structure in the unfolded states at either pH 2.7 and 3.0 M urea or pH 7.6 and 8.0 M urea.

After the initiation of refolding, two phenomena were observed. First, several residues were protected after only 8 ms. Second, the degree of protection was different for different residues. The simplest explanation of these observations is that a single intermediate is formed and that different residues in this intermediate display different exchange behaviors. Although the data are consistent with a single pathway for refolding T4 lysozyme under these conditions, we cannot rigorously rule out the possibility of multiple pathways and multiple intermediate structures in the path from the unfolded to folded state. For example, during the pulse-labeling period at high pH, it is possible that the distribution among a population of intermediates might change. Thus the observed proton occupancy might reflect the accessibility of one subpopulation and a pH-dependent shift in the amount of that population rather than the accessibility of a single homogeneous intermediate. Such a multiple intermediate mechanism for folding seems unlikely because the overall rate of folding is pH independent from pH 5 to pH 9 (Lu and Dahlquist, in preparation). This would appear to rule out a gross change in the amount and kinetic competence of the imagined intermediates as a function of pH. These issues will be further addressed in future work.

We cannot be certain that the intermediate(s) we have detected is (are) kinetically competent. However, preliminary results suggest that dramatic slowing of folding can occur in mutants in which residues in some protected regions of the intermediate(s) are changed (unpublished observations). This argues that the observed intermediate(s) may play an important role in the folding pathway.

The pH dependence of the proton occupancy at different sites indicates that exchange probably occurs through an EX2 mechanism (Hvidt & Nielsen, 1966; Englander & Kallenbach, 1984; Udgaonkar & Baldwin, 1990). In this mechanism, there is a rapid preequilibrium to an exchange-competent state followed by base-catalyzed exchange of the amide. The fast preequilibrium determines the amount of the competent state available for exchange. There are several possible interpretations for the nature of the preequilibrium step including (1) global unfolding of the intermediate(s), (2) a partial unfolding of the region of interest, or (3) the penetration of solvent to the sites of interest. For any of these interpretations, the protection factor observed for any site in an intermediate is a measure of the minimum stability of the intermediate itself. The calculated protection factors of 20 or more for each of the three regions argue that the intermediate(s) is (are) rather stable and that most if not all the molecules form this intermediate or intermediates as they are converted from the unfolded to the folded state.

Our pulsed hydrogen-exchange experiments done at different pH values show that in the intermediate(s), several residues are substantially protected from solvent exchange. Still, any given residue in the intermediate(s) is much more susceptible to exchange than it is in the final native structure. Most of the protected residues in β-sheets in the N-terminal domain and helix A and helix E in the C-terminal domain are protected by 20–180-fold relative to the free amide. Several residues in helix C are also about 20-fold protected relative to the free amide. Protection factors of similar magnitude were observed in the partly folded apomyoglobin intermediate (Hughson et al., 1990). Such protection factors are consistent with the formation of somewhat isolated regions of secondary structure. All the remaining probe residues are less than 10-fold protected. This low level of protection suggests that these residues may not have formed any hydrogen-bonded structures in the intermediate(s), and it is possible that these residues are still in rapidly fluctuating random coil-like disordered configurations.

Hydrogen–deuterium exchange of the native structure studied by NMR has shown that amide protons of all 84 residues used as probes undergo very slow exchange (McIntosh et al., in preparation) and are protected by at least a factor of 10^4. The region of the protein that shows the highest degree of protection from hydrogen exchange corresponds to residues 93–105. These correspond to helix E in the native structure of the protein. The crystal structure of the native state shows that this helix is the most buried (Matthews & Remington, 1974). It is reasonable for this helix to be formed earliest during refolding, and this region is the most protected in the intermediate(s). Helix A is in close contact with helix E, and helix A is also protected in the intermediate(s). It is possible that these protected regions of the intermediate(s) form a framework for refolding of the rest of the protein. These two regions of the intermediate(s) may have α-helical secondary structure similar to those found in native T4 lysozyme. The protection observed in the N-terminal domain of the intermediate(s) has a pattern of protection that is suggestive of β-sheet structure, corresponding to the structure of this region in the final folded state. Additional experiments will be needed to more exactly establish the likely structures of these regions in the intermediate(s).

Previous kinetic circular dichroism studies have indicated that secondary structure is formed before the final tertiary structure during refolding of α-lactalbumin (Gilmanshin & Ptitsyn, 1987; Kuwajima et al., 1985), hen egg white lysozyme (Ikeguchi et al., 1985), carbonic anhydrase B (McCoy et al., 1980; Dolgikh et al., 1984; Semistonov et al., 1987), and ferrocytochrome c and β-lactoglobulin (Kuwajima et al., 1987). NMR trapping techniques applied to RNase A, horse heart cytochrome c, barnase, and hen egg white lysozyme have detected the presence of intermediates with a subset of native secondary structure (Udgaonkar & Baldwin, 1988; Roder et al., 1988; Bycroft et al., 1990; Miranker et al., 1991). It appears that T4 lysozyme also follows this trend. Our data suggest that the bulk of the lysozyme molecules form a very nativelike structure during the 170-ms relaxation process. After this time all the amides used as probes are strongly protected. The three tryptophan residues used to follow the folding time course by fluorescence are all found in the C-terminal domain. Unfortunately, the amide protons of all three (residues 126, 138, and 158) exchange fast in the native protein and are unsuitable to be used as probes. Since we observe no early changes in fluorescence, it appears that the environments around the indole side chains are not nativelike at early times. These observations suggest that the early intermediate we have observed has not formed all the side-chain interactions characteristic of the native state.

Models for protein folding have been recently reviewed by Kim and Baldwin (1982, 1990), Harrison and Durban (1985), Ptitsyn (1987), Jaenicke (1988), and Baldwin (1989). Current models include the framework model (Kim & Baldwin, 1982), the subdomain model (Oas & Kim, 1988; Staley & Kim, 1990), and the molten globule model (Ptitsyn, 1987; Kuwajima, 1989; Ptitsyn et al., 1990). In the framework model, an intermediate with native secondary structure is formed before the tertiary structure is locked in place. In the subdomain model, the protein folds in parts and different subdomains coalesce later. In the molten globule model, a compact intermediate containing a high content of secondary structure is formed early during refolding. There are features

of both the framework and molten globule models in the refolding of T4 lysozyme. Secondary structure is formed at some regions before much tertiary structure has been formed and before the interactions between the N-terminal and C-terminal domains have been completed. The data do not appear to support the subdomain model even though T4 lysozyme has two relatively distinct domains.

Hen egg white (HEW) lysozyme is also a two-domain protein. Its folding pathway has been studied by competition hydrogen exchange and 2-D NMR. Refolding of the two domains appears to be sequential (Miranker et al., 1991). In contrast to HEW lysozyme, folding of the two domains of T4 lysozyme appears to be concerted. There are significant structural similarities in the crystal structures of HEW and T4 lysozymes, but there is essentially no similarity in amino acid sequence (Weaver et al., 1985). The β-sheets and several helical regions of the two proteins are homologous in structure. Interestingly, the patterns of amide protection during folding are quite different. For example, the β-sheet region of T4 lysozyme forms rapidly while the homologous β-sheet of HEW lysozyme shows no evidence of early structure. Only one region of structural similarity (helix A of T4 and helix B of HEW) shows evidence of early structure formation in both proteins. This suggests that amino acid sequence rather than the secondary structure of the folded state may control early events in folding.

Future work from our laboratory will attempt to elucidate the kinetic role of the intermediate(s) and the possible interactions between the domains during the folding process.

ACKNOWLEDGMENTS

We thank Lawrence McIntosh, Eric Anderson, Michael Strain, Amy Roth, and Andy Morton for help and advice; Hale Nicholson for providing single proline mutants; Michael Schimerlik for the use of his stopped-flow instrument (supported by a grant from Environmental Health Services, EF00210); Mark Young for help in using rapid-mixing quench-flow apparatus; and Dennis Hare for supplying the FTNMR software.

REFERENCES

Alber, T. A., Dao-pin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987) *Biochemistry 26*, 3754–3758.

Anderson, D. E., Becktel, W. J., & Dahlquist, F. W. (1990) *Biochemistry 29*, 2403–2408.

Baldwin, R. L. (1989) *Trends Biochem. Sci. (Pers. Ed.) 14*, 291–294.

Baldwin, R. L., & Roder, H. (1991) *Curr. Biol. 1* (No. 4), 218–220.

Becktel, W. J., & Baase, W. A. (1987) *Biopolymers 26*, 619–623.

Bell, J. A., Dao-pin, S., Faber, R., Jacobsen, R., Karpusas, M., Matsumura, M., Nicholson, H., Pjura, P. E., Tronrud, D. E., Weaver, L. H., Wilson, K. P., Wozniak, J. A., Zhang, X.-J., Alber, T., & Matthews, B. W. (1990) *Use of X-ray Crystallography in the Design of Antiviral Agents*, pp 233–245, Academic Press, NY.

Bycroft, M., Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. R. (1990) *Nature (London) 346*, 488–490.

Chen, B., Baase, W. A., Nicholson, H., & Schellman, J. A. (1992) *Biochemistry 31*, 1464–1476.

Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., & Ptitsyn, O. B. (1984) *FEBS Lett. 165*, 88–92.

Elöve, G. A., & Roder, H. (1991) in *Protein Folding* (Georgiu, G., Ed.) ACS Symposium Series 70, pp 50–63, Washington, DC.

Englander, W. S., & Kallenbach, N. (1984) *Q. Rev. Biophys. 16*, 521–655.

Gilmanshin, R. I., & Ptitsyn, O. B. (1987) *FEBS Lett. 223*, 327–329.

Harrison, S. C., & Durbin, R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4028–4030.

Hvidt, A., & Nielsen, S. O. (1966) *Adv. Protein Chem. 21*, 287–386.

Hudson, B., Harris, D., Gray, T., & McIntosh, L. (1987) *Biophys. J. 51*, 419a.

Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544–1548.

Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6965–6972.

Jaenicke, R. (1988) *Prog. Biophys. Mol. Biol. 49*, 117–237.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–189.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem.*, 631–660.

Kuwajima, K. (1989) *Proteins 6*, 87–103.

Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.

Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S., & Nagamura, T. (1987) *FEBS Lett. 221*, 115–118.

Loftus, D., Gbenle, G. O., Kim, P., & Baldwin, R. L. (1986) *Biochemistry 25*, 1428–1436.

Matsumura, M., & Matthews, B. W. (1989) *Science 243*, 792–794.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.

Matthews, B. W., & Remington, S. J. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4178–4182.

McCoy, L. F., Rowe, E. S., & Wong, K. P. (1980) *Biochemistry 19*, 4738–4743.

McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., & Dahlquist, F. W. (1990) *Biochemistry 29*, 6341–6362.

Miranker, A., Radford, S. E., Karplus, M., & Dobson, C. M. (1991) *Nature (London) 349*, 633–636.

Molday, R. S., Englander, S. W., & Kallen, R. G. (1972) *Biochemistry 11*, 150–159.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.

Oas, T. G., & Kim, P. S. (1988) *Nature (London) 336*, 42–48.

Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.

Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20–24.

Robertson, A. D., & Baldwin, R. L. (1991) *Biochemistry 30*, 9907–9914.

Roder, H., Elöve, G. A., & England, W. S. (1988) *Nature (London) 335*, 700–704.

Semisotnov, G. N., Rodionova, N. A., Kutyshenko, V. P., Ebert, B., Blanck, J., & Ptitsyn, O. B. (1987) *FEBS Lett. 224*, 98–113.

Staley, J. S., & Kim, P. S. (1990) *Nature (London) 334*, 685–688.

Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature (London) 335*, 694–699.

Udgaonkar, J. B., & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8197–8201.

Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189–199.

Weaver, L. H., Grütter, M. G., Remington, S. J., Gray, T. M., Isaacs, N. W., & Matthews, B. W. (1985) *J. Mol. Evol. 21*, 97–111.