Tertiary Interactions in the Folding Pathway of Hen Lysozyme: Kinetic Studies Using Fluorescent Probes†

Laura S. Itzhaki,‡,§ Philip A. Evans,*,‡ Christopher M. Dobson,‖ and Sheena E. Radford‖

Cambridge Centre for Molecular Recognition and Department of Biochemistry, Cambridge University, Tennis Court Road, Cambridge CB2 1QW, U.K., and Oxford Centre for Molecular Sciences and New Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QT, U.K.

Received September 30, 1993; Revised Manuscript Received February 8, 1994®

ABSTRACT: The refolding kinetics of hen lysozyme have been studied using a range of fluorescent probes. These experiments have provided new insight into the nature of intermediates detected in our recent hydrogen-exchange labeling studies [Radford, S. E., *et al.* (1992) *Nature 358*, 302–307], which were performed under the same conditions. Protection from exchange results primarily from the development of stabilizing side-chain interactions, and the fluorescence studies reported here have provided a new perspective on this aspect of the refolding process. The intrinsic fluorescence of the six tryptophan residues and its susceptibility to quenching by iodide have been used to monitor the development of hydrophobic structure, and these studies have been complemented by experiments involving binding to a fluorescent hydrophobic dye 1-anilino-naphthalenesulfonic acid (ANS). Formation of fixed tertiary interactions of aromatic residues has been monitored by near-UV circular dichroism, while development of a competent active site has been probed by binding to a competitive inhibitor bearing a fluorescent label, 4-methylumbelliferyl-*N*,*N*′-diacetyl-β-chitobiose. The combination of these techniques has enabled us to monitor the development both of the hydrophobic core of the protein and of interactions between the two folding domains. If the behavior of the tryptophans is representative of the hydrophobic residues of the protein in general, it seems that collapse is already substantial in species formed within the first few milliseconds of refolding and is highly developed in later intermediates which nonetheless appear to lack many fixed tertiary interactions. Some of the details of the native structure, including the active site which is formed at the interface between the folding domains, develop only in the slowest stages of folding, even though in a subset of molecules stable native-like structure exists in both domains at an early stage of folding. These late events probably involve fine adjustments of side-chain packing and formation of specific ionic interactions that occur in the native state.

Investigation of the pathways by which globular proteins fold is a key approach to understanding their structural organization. Much evidence suggests that folding proceeds, at least in most cases, via transient partially ordered intermediates, so that pathways can be characterized in terms of the structural, kinetic, and thermodynamic properties of these species (Kim & Baldwin, 1990; Matthews, 1991; Creighton, 1993). A range of chemical and spectroscopic methods, including stopped-flow optical methods, hydrogen-exchange labeling (Baldwin & Roder, 1992; Englander & Mayne, 1992; Baldwin, 1993), and protein engineering techniques (Jennings *et al.*, 1992; Fersht & Serrano, 1993), have provided key information about different aspects of folding pathways. The ephemeral nature of intermediates and the extreme rapidity of at least a number of the steps occurring during folding make the acquisition of the detailed information required to describe folding at the molecular level a demanding problem. The information provided by these different techniques is complementary, and thus, by utilizing a combination of methods to study the same system under identical conditions, a much fuller picture of the events occurring during folding can be drawn than is possible from these methods used in isolation.

We have applied this general approach to study the folding of hen lysozyme. This small monomeric protein contains many of the structural features commonly found in globular proteins, including α- and 3^{10}-helices, β-sheets, loops, and turns (Figure 1). Pulsed hydrogen-exchange labeling experiments have revealed that, under the conditions used, the folding process involves a number of well-defined steps and population of specific intermediates (Radford *et al.*, 1992). Specifically, there are two distinct folding domains, so that the α-helical part of the native structure folds, at least in the majority of molecules, more rapidly than a second domain, which includes a triple-stranded β-sheet, a 3^{10}-helix, and a long loop (Figure 1). Further, neither folding domain is stabilized in a single kinetic step: two phases of comparable amplitude (*ca.* 40%) with average time constants close to 5 and 65 ms can be resolved for residues in the α-domain, while time constants of approximately 10 and 340 ms are characteristic of amides in the β-domain, the fast phase in this case accounting for only about 25% of the total amplitude. This kinetic diversity reflects the existence of parallel folding pathways which presumably are a consequence of conformational heterogeneity at a crucial, very early stage of refolding. All of these studies have been performed under conditions in which the four disulfide cross-links of the protein remain intact, and one possibility is that slow kinetic phases arise from populations with one or more disulfides in an inappropriate orientation (Chaffotte *et al.*,

† We acknowledge support by the Royal Society and by the SERC and MRC through the Oxford Centre for Molecular Sciences and the Cambridge Centre for Molecular Recognition. S.E.R. is a Royal Society 1983 University Research Fellow. C.M.D. is an International Research Scholar of the Howard Hughes Medical Institute.

* Author to whom correspondence should be addressed.
‡ Cambridge University.
§ Present address: University Chemical Laboratories, Cambridge University, Lensfield Road, Cambridge CB2 1EW, U.K.
‖ Oxford University.

® Abstract published in *Advance ACS Abstracts,* March 15, 1994.

0006-2960/94/0433-5212$04.50/0

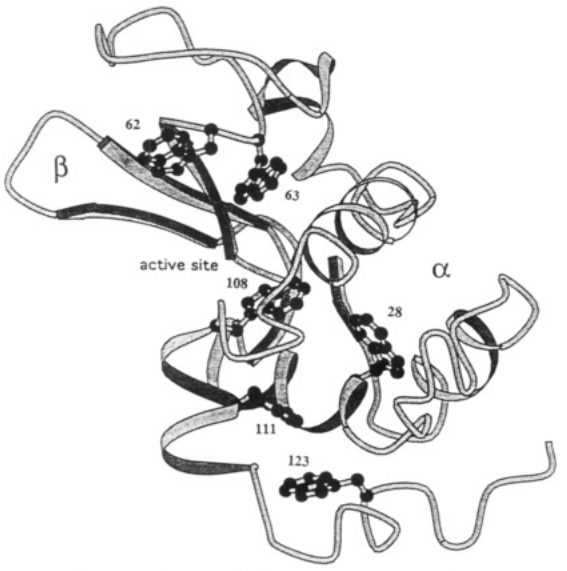

FIGURE 1: Schematic view of the native structure of hen lysozyme, including the distribution of the six tryptophan residues. The two folding domains identified by hydrogen-exchange labeling (Radford *et al.*, 1992) correspond closely to the two structural lobes that can readily be distinguished, one formed from four α-helices (labeled α) and the second (labeled β) including the triple-stranded β-sheet and the adjacent 3^{10}-helix. The active site is formed at the interface between these lobes, as indicated. The drawing was generated using the program MOLSCRIPT (Kraulis, 1991), from crystallographic coordinates (Balke *et al.*, 1965).

1992). An additional folding pathway with a much longer time scale ($\tau \sim 16$ s) occurs in about 20% of molecules and is probably limited by *cis*/*trans* isomerism of the two proline residues in the amino acid sequence (Kato *et al.*, 1981).

Kinetic CD measurements in the far-UV have been particularly informative about the very early steps in the folding of lysozyme (Radford *et al.*, 1992; Chaffotte *et al.*, 1992). The results indicate that substantial secondary structure is formed much more quickly than the persistent interactions detected through hydrogen-exchange protection. The subsequent time course is complex, with the ellipticity at 225 nm passing through a minimum such that its magnitude actually exceeds the value of the native state before recovering more slowly to its final value. A detailed study has suggested that a complement of secondary structure comparable in extent to that of the native state is acquired within the dead time (4 ms) of the experiment while subsequent excursions seem likely to reflect the changing contribution of non-peptide chromophores, probably disulfide bonds, to the CD spectrum (Chaffotte *et al.*, 1992). In contrast to the rapid development of secondary structure reflected in the far-UV, CD measurements in the near-UV have suggested that ordering of aromatic residues in fixed orientations occurs much more slowly, on a time scale similar to that of the major protection phase of residues in the β-sheet.

Despite the wealth of detail derived from these studies, there remain several key aspects of the folding pathway of lysozyme about which relatively little is understood. These include the degree to which secondary structure formation is accompanied by development of hydrophobic interactions in early intermediates and the time scale of formation of the fully native, functionally competent enzyme. To address these issues, we have undertaken a further series of stopped-flow experiments, under the same conditions used for our earlier pulse-labeling and CD experiments (Radford *et al.*, 1992), using various extrinsic and intrinsic fluorescent probes. First, the development of a functional active site was monitored through its binding of a fluorescent substrate analogue, 4-methylumbelliferyl-*N*,*N*′-diacetyl-β-chitobiose (MeU-di-NAG). This is a particularly interesting aspect of the folding process, as it entails bringing together residues from the two structural domains found to be formed noncooperatively in the pulse-labeling experiments.

In a second series of experiments the intrinsic fluorescence of lysozyme was examined. This is dominated by the six tryptophan residues, and the changing intensity of this emission was readily monitored during the refolding process. One of the major values of intrinsic fluorescence in studies of proteins is its sensitivity to the solvent exposure of the side chains involved, and this has led to its extensive use in folding studies. Some proteins contain tryptophan residues which provide a convenient probe; in other instances these have been introduced by protein engineering (Smith *et al.*, 1991; Khorasanizadeh *et al.*, 1993). However, the most readily followed parameter, the total intensity, cannot be interpreted straightforwardly because this is dependent on a variety of factors (Eftink, 1991) and any of these could contribute to its variation during folding. An additional series of experiments involving the quenching of fluorescence by iodide was undertaken, therefore, to monitor more directly the exclusion of indole side chains from solvent and hence the buildup of hydrophobic clusters in the folding process.

Finally, a different perspective on the time development of hydrophobic collapse was gained from studies of the interaction of the refolding protein with 1-anilinonaphthalenesulfonic acid (ANS). The relatively loose packing of "molten globule" states apparently favors binding of this hydrophobic dye molecule, and the consequent enhancement in its fluorescence quantum yield has now been used extensively for detecting these partially folded species (Ptitsyn *et al.*, 1990; Semisotnov *et al.*, 1991). Combining the results of all these approaches has provided new insight into the nature of the various folding intermediates, and together with the results of the hydrogen-exchange labeling and CD experiments, this has enabled us to construct a more detailed model for the folding pathway of the enzyme.

MATERIALS AND METHODS

Hen egg white lysozyme (grade I), ANS, *N*-acetyltryptophanamide, and Gly-Gly-Trp-Ala were purchased from Sigma. MeU-diNAG was from Fluka Chemical Co. The concentration of MeU-diNAG was determined using a molar extinction coefficient of 1.24×10^4 M^{-1} cm^{-1} at 316 nm (van Landshoot *et al.*, 1977).

Stopped-Flow Measurements. Fluorescence detection of the folding kinetics of hen lysozyme was achieved in most experiments using a Biologic SFM3 stopped-flow module, coupled via a fiber-optic to an external fluorometer (Jasco FP777) which served as the light source. An excitation bandwidth of 10 nm was employed, and emission was measured above a chosen wavelength by inserting an appropriate colored glass filter between the stopped-flow cell and the photomultiplier tube. A Biologic PMS200 detection system was used. The temperature in the stopped-flow unit was maintained at 20 °C using water circulating from a thermostated bath. Under the experimental conditions employed (flow rate 5.5 or 7.9 mL/s) the dead time of the equipment was 3 or 2 ms, respectively. The inhibitor binding and fluorescence quenching data were repeated under identical refolding conditions using an Applied Photophysics stopped-flow system. This gave enhanced sensitivity, and the comparison of data obtained on the different systems eliminated the possibility of misinterpretation of instrumental artifacts.

Stopped-flow CD measurements were performed using either a Jasco J720 spectropolarimeter or a Jobin-Yvon CD6

circular dichrograph, coupled in either case to a Biologic SFM3 stopped-flow module. In all experiments unfolded lysozyme (20 mg/mL in 6 M guanidine hydrochloride) was diluted 11-fold with 20 mM sodium acetate buffer, pH 5.2, giving a final protein concentration of 125 μM. Refolding was typically followed for up to 2 s, using a sampling time of 1 or 2 ms. Control experiments were performed in which 6 M guanidinium chloride solution containing camphorsulfonic acid or oxidized ribonuclease A (which does not undergo a conformational charge upon dilution from 6 M guanidine hydrochloride) was mixed with the refolding buffer; in all cases any changes observed were small compared to those resulting from refolding of the protein and were therefore ignored.

Intrinsic Fluorescence. An excitation wavelength of 280 nm was used, and total fluorescence above 310 nm was monitored using a suitable colored glass long-band filter. The change in intrinsic fluorescence was monitored directly with time after the syringes had stopped. Data were corrected for contributions from scattered light by subtracting the baseline time course. The dead time in these experiments was 3 ms. The ionic strength dependence of the refolding kinetics was investigated by inclusion of various concentrations of NaCl in the refolding buffer.

Fluorescence Quenching by Iodide. For these experiments the excitation wavelength was 280 nm and total fluorescence emission above 310 nm was measured as described above. The experiments were performed in two different ways. In the first series, the protein (20 mg/mL in 6 M guanidine hydrochloride) was diluted 11-fold with 20 mM sodium acetate, pH 5.3, containing different concentrations of NaI and NaCl, such that a constant ionic strength was maintained; the progress curves were then monitored after the syringes had stopped moving (dead time ~2 ms). To avoid any kinetic complications arising from the different ionic compositions during refolding in this experiment, an alternative strategy was also developed. In this series of experiments the enzyme was diluted initially with 20 mM sodium acetate, pH 5.2. After a variable refolding time (2 ms–2 s), a second mix was performed with an equal volume of 0.6 M NaI or 0.6 M NaCl. The fluorescence intensity was then determined as quickly as possible after this second mix, the only delay being the dead time (2 ms) inherent in the apparatus. In many cases it was possible to measure the fluorescence of a constant-age sample during continuous flow through the cell, thus permitting data averaging. The extent of quenching as a function of refolding time was investigated by varying the delay between the mixes. This method is clearly much more cumbersome than the direct inclusion of quencher in the refolding buffer, but it ensures that the refolding conditions in different experiments are identical.

Binding of ANS. An excitation wavelength of 390 nm was used, the total fluorescence emission above 435 nm was monitored. Experiments were performed both with ANS included in the refolding buffer and by addition of ANS in a second mix after various refolding times, as described above for iodide quenching. For each experiment the final concentrations of protein and ANS were 61 and 90 μM, respectively.

Binding of MeU-diNAG. Fluorescent emission was measured above 338 nm using an excitation wavelength of 330 nm. Experiments were performed by an 11-fold dilution of 20 mg/mL lysozyme in 6 M guanidine hydrochloride with a solution of 20 mM sodium acetate, pH 5.2, containing 6 μM MeU-diNAG. For each experiment a relevant baseline was acquired immediately prior to the refolding kinetics by diluting 6 M guanidine hydrochloride with 10 volumes of the refolding buffer under identical conditions. In this case a double-mix approach was not used because it proved difficult to construct a reproducible refolding curve from the relatively small fluorescence changes observed. The rate of binding of the fluorescence inhibitor to the native enzyme was determined, under conditions identical to the final refolding conditions, by mixing a solution of native lysozyme (20 mg/mL) with 10 volumes of the refolding buffer, containing 6 μM MeU-diNAG and 0.6 M guanidine hydrochloride. Control experiments to confirm that the MeU-diNAG was binding specifically to the active site were performed by diluting the refolding enzyme into the same buffer mixture containing, in addition, 120 μM triNAG. Under these conditions the triNAG effectively saturates the active site (Imoto *et al.*, 1972).

Kinetic Analysis. Kinetic parameters were obtained by nonlinear least-squares fitting to one or two exponential phases, using the program Kaleidagraph. The amplitudes for the individual phases were normalized with respect to the total fluorescence change upon refolding.

RESULTS

Kinetics of Binding of the Fluorescent Inhibitor MeU-diNAG. Binding of a competitive inhibitor during refolding provides a powerful way to determine directly the rate of appearance of the native state of an enzyme. In lysozyme the active site is formed at the interface between the two folding domains (Figure 1), and the kinetics of its organization, therefore, provide particular insight into the process by which the two domains become integrated into a cooperative native structure. MeU-diNAG is a substrate analogue which binds in the active site of lysozyme to form a 1:1 complex, with a consequent increase in its fluorescence that can conveniently be monitored in a stopped-flow experiment (Yang & Hamaguchi, 1980).

Measurements of the kinetics of binding of MeU-diNAG to native lysozyme, under the conditions of the refolding experiments, revealed an exponential process with a time constant of about 5 ms. In refolding experiments, therefore, the rates of any phases found to be slower than this must be limited by events in the folding process. To make detection of inhibitor binding during refolding experimentally tractable, it was convenient to include the inhibitor in the refolding buffer and to perform a straightforward single-mix stopped-flow experiment. It was then necessary, however, to ensure that the presence of the inhibitor did not affect the refolding kinetics. This was investigated by monitoring the change in intrinsic fluorescence of the tryptophan residues of the protein as it folds (see below). This experiment confirmed that the presence of the MeU-diNAG had no observable effect on the folding kinetics, consistent with previous results which suggested that the stabilizing effect of such inhibitors is manifested entirely in the rate of unfolding (Segawa & Sugihara, 1984).

The kinetic progress curve of the binding of MeU-diNAG to lysozyme as it refolds is shown in Figure 2a. The data fit very well to a single exponential with a time constant of around 360 ms. The fitted curve accounts for the entire expected kinetic amplitude, based on the measured fluorescence of the free inhibitor and of the enzyme-bound complex under identical conditions. To confirm that the inhibitor binds specifically to the active site of the refolding enzyme, a control experiment was performed in which a 20-fold molar excess, relative to the MeU-diNAG concentration, of the inhibitor triNAG was also included in the refolding buffer. Under these conditions the change in fluorescence of MeU-diNAG during folding was almost completely abolished, presumably because the strongly binding, nonfluorescent triNAG inhibitor effectively saturates the active site.

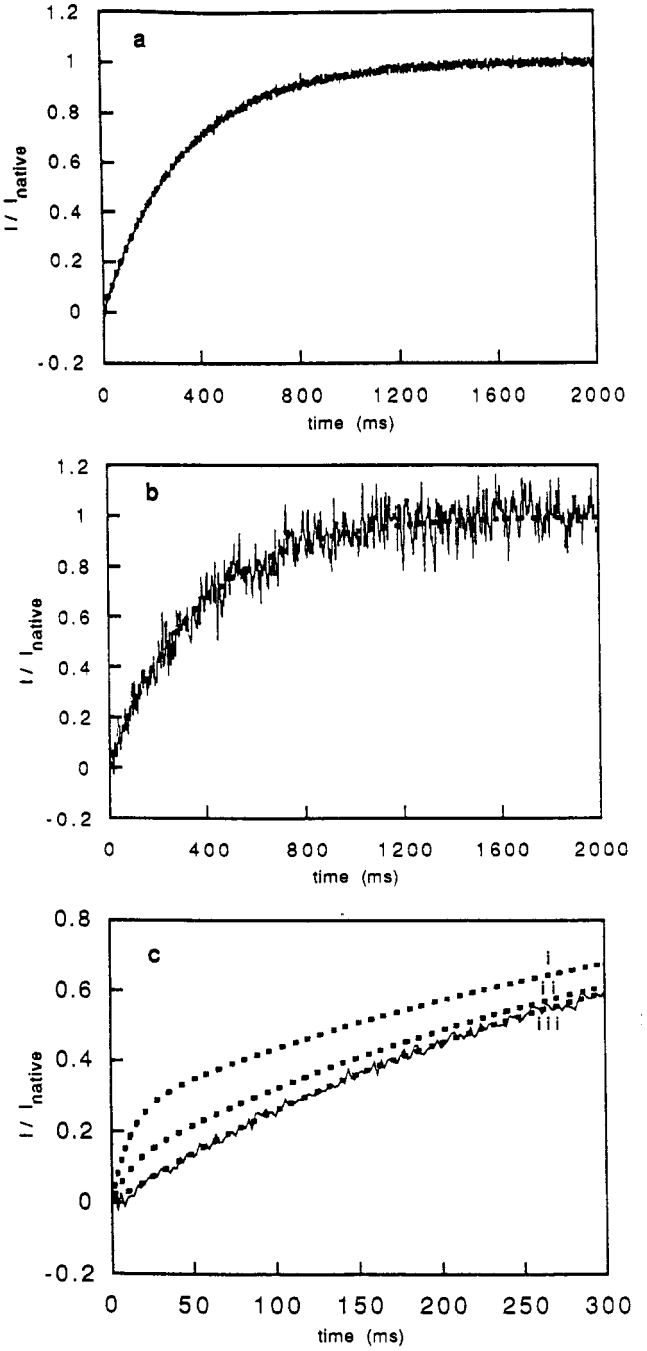

FIGURE 2: Time course of refolding of lysozyme in 0.54 M Gdn-HCl, pH 5.2, 20 °C, monitored by (a) the intensity (I) of fluorescence of MeU-diNAG included in the refolding buffer and (b) circular dichroism at 289 nm. In both cases the data have been fitted to a single exponential (dotted curves) with time constants of 360 and 330 ms, respectively. The fast phase in the near-UV CD reported in our earlier paper appears to have been artifactual (Radford *et al.*, 1992). (c) Comparison of the best fitting single exponential (iii) to the early part of the MeU-diNAG trace, with simulated curves with the same slow time constant (360 ms) but including a fast phase of time constant 11 ms and amplitudes of 25% (i) or 10% (ii), which might have been consistent with the hydrogen-exchange labeling data (see text).

Figure 2b shows the kinetics of the development of the near-UV CD during the refolding reaction. The curve fits well to a single exponential with a time constant close to 330 ms and accounts for the entire kinetic amplitude based on the measured ellipticity of the native and denatured states. This rate is similar to that of active site formation, as reflected by MeU-diNAG binding, and also that of the slowest event detected in the hydrogen-exchange labeling studies which is associated with protection of amides in the β-domain of the majority of molecules (Radford *et al.*, 1992). Thus, it is apparent that in the refolding of most molecules, at least, formation of a persistent β-domain is synchronous with formation of the fully native molecule with a functional active site.

Not all of the refolding molecules, however, acquire protection from exchange in the β-domain on this slow time scale. Thus, in the rapid protection phase ($\tau \sim 5$–10 ms) around 25% of amides in the β-domain and around 40% of amides in the α-domain become protected from exchange. Analysis of the populations of these protected species by electrospray mass spectrometry has revealed that a species protected in both the α- and β-domains occurs in about 20% of molecules on this time scale (Miranker *et al.*, 1993). The simplest interpretation of this result by itself would have been that these molecules have adopted the native structure. If this were the case, however, a corresponding fast phase should be observed in both the inhibitor binding and near-UV CD experiments. The present experiments show clearly that there is no evidence for such a fast phase in the MeU–diNAG binding curve; the excellent sensitivity of the experiment permits us to conclude confidently that a single exponential provides an entirely satisfactory fit (Figure 2c). Similarly, a single-exponential function gave a better fit to the near-UV CD data than did simulated double exponentials. Thus, it is clear that despite the rapid protection of a native complement of amides in this phase, the refolding molecules are clearly not yet fully native but represent a further intermediate which is significantly populated only on a minor folding pathway. This species is more highly structured than the intermediates exhibiting protection only in the α-domain, but its aromatic residues are still not fixed in an asymmetric environment and the docking of the two folding domains is not yet fully established.

Changes in Intrinsic Tryptophan Fluorescence on Refolding. The kinetics of folding of lysozyme as reflected in the intrinsic fluorescence of its tryptophan residues are shown in Figure 3a. The profile is complex and is characterized by at least three kinetic phases. First, in the dead time of the experiment a substantial decrease in intensity occurs, so that at the earliest measurable time its magnitude is reduced by a factor of about 1.9 relative to that of the denatured enzyme in 6 M guanidine hydrochloride. In part this can be attributed to a solvent effect, since fluorescence intensity is observed to vary with guanidine hydrochloride concentration even in the case of the small peptide Gly-Gly-Trp-Ala, the fluorescence intensity of which was observed to decrease by a factor of approximately 1.6 when the guanidinium chloride concentration was reduced from 6 to 0.54 M. This factor was used to estimate the fluorescence intensity of fully unfolded lysozyme in 0.54 M guanidine hydrochloride. The value is still larger by a factor of around 1.2 than that observed at the first measurable time of refolding in the stopped-flow experiment (Figure 3a). This suggests, therefore, that there are significant changes in fluorescence quantum yields during these first 2 ms of refolding.

The second phase of the curve comprises a further decrease in fluorescence, so that by around 100 ms the intensity reaches a minimum, at a value around 60% of that of the native enzyme. The MeU-diNAG binding data show that some 30% of molecules have actually been restored to the native state by this time, thus accounting for about half of this emission. The remaining 70% of molecules, which are in partially folded states, contribute the other half of the fluorescence and therefore we can deduce that the total quantum yield in these molecules is only about 40% of its native-state value, suggesting that there are effective, nonnative intramolecular quenching mechanisms operative in the partially folded state. In the

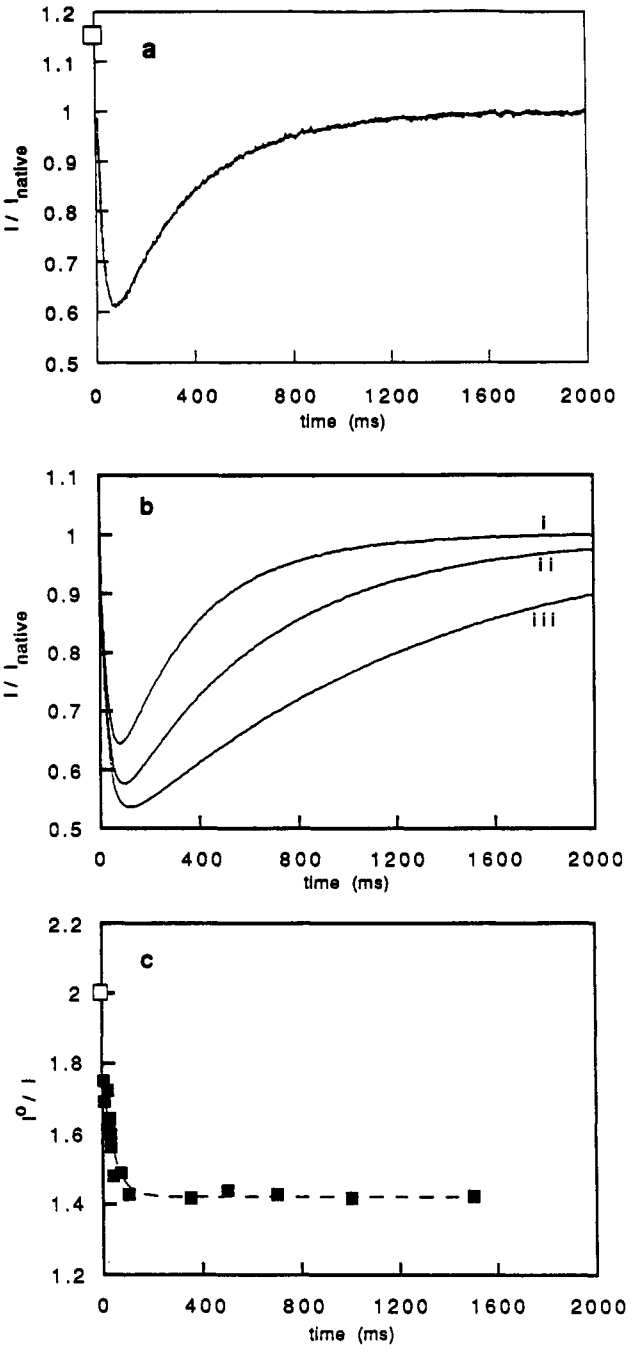

FIGURE 3: (a) Time course of refolding of lysozyme in 0.54 M Gdn-HCl, pH 5.2, 20 °C, monitored by intrinsic fluorescence excited at 280 nm and monitored as total intensity above 320 nm (I). The curve has been fitted to a sum of two exponentials, with time constants 25 and 340 ms (dotted line). Also shown (□) is the predicted fluorescence intensity of fully unfolded lysozyme under the refolding conditions (see text). The amplitudes of the kinetic phases relative to the total change from this nominal starting state to the native state (observed at 2 s) are +120% (dead time), +290% (25 ms), and −310% (340 ms). (b) Effect of salt concentration on the kinetics of refolding of lysozyme monitored by intrinsic fluorescence of tryptophan residues. The added NaCl concentrations were 0 (i), 1.5 (ii), and 3.0 M (iii). (c) Quenching of intrinsic fluorescence by 0.3 M iodide at various times during the refolding of lysozyme. The conditions of the experiment were as described above. The quenching is expressed as the ratio of the intensity when a 1:1 dilution of the refolding mixture with 0.6 M KCl was performed (I^0) to that observed when the partially folded protein was diluted with 0.6 M KI (I). The time course has been fitted to a single exponential, with time constant 40 ms (dotted curve). Also shown (□) is the predicted fluorescence quenching for a fully unfolded state under these conditions (see text).

slowest phase the fluorescence intensity then increases once again, to its limiting value. These two kinetic phases were

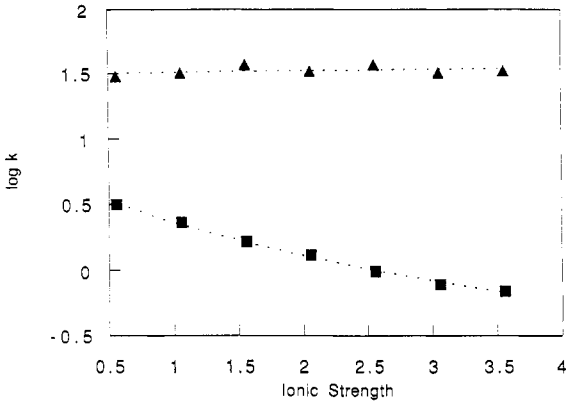

FIGURE 4: Ionic strength dependence of the rates of the fast (▲) and slow (■) phases of refolding of lysozyme as detected by intrinsic fluorescence (see Figure 3b). The ionic strengths indicated are calculated from the concentrations of all the salts present under the final refolding conditions, including the guanidine hydrochloride, the acetate buffer, and various amounts of additional NaCl. The negatively sloping plot obtained for the slow phase is consistent with formation of interactions between opposing charges in the transition state. The fast phase is virtually independent of the ionic strength. The best fit curves (dotted lines) are to the relationship $\log k = B + AI^{1/2}$, where A and B are constants and I is the ionic strength. This derives from the Debye–Hückel limiting law and characterizes the classical primary kinetic salt effect. The behavior of the slow phase appears to fit surprisingly well to this model, given the very high ionic strengths involved and the fact that this is clearly not a bimolecular reaction in free solution.

fitted to the sum of two exponentials, the time constants for which are 25 ms for the intensity decrease and 340 ms for the recovery. These data show that significant changes in the environments of some or all of the six tryptophan residues occur during the dead time of refolding, in forming subsequent folding intermediates, and again on reorganization into the native state. The nature of the conformational changes involved cannot be deduced, however, because a wide variety of factors can influence fluorescence intensity. To gain more insight, in particular into the role of changes in hydration of the tryptophan side chains, the susceptibility of the intrinsic fluorescence to quenching by iodide was investigated.

Quenching of Fluorescence by Added Solutes. Exclusion of a charged quenching agent such as iodide ion from the protein interior can potentially provide a more direct measure of the exclusion of tryptophan residues from solvent as the protein folds (Lehrer, 1971; Garvey *et al.*, 1989), presumably reflecting development of the hydrophobic core. To apply this approach to lysozyme, the influence of added salts on the refolding kinetics was first investigated by monitoring the tryptophan fluorescence during refolding in the presence of different concentrations of NaCl (0–3 M) (Figure 3b). In the presence of the added salt the fluorescence changes occurred qualitatively in the same three kinetic phases observed in the absence of additional salts, as described above. However, while the kinetic amplitudes of the two resolved phases and the rate of the faster phase were intensitive to the added salt, the rate of the slowest phase of folding was significantly diminished as the salt concentration was increased. Qualitatively similar results were obtained when the effect of NaI on the refolding rate was investigated. The ionic strength dependence of the rates is shown in Figure 4. The behavior of the slow phase is indicative of a kinetic salt effect, the most obvious explanation being that the formation of specific favorable ionic interactions within the protein structure occurs only in the final, slow phase of folding. This is further evidence in support of the conclusion that the details of the tertiary structure are not fixed until this late stage in folding.

The sensitivity of the refolding kinetics to the final refolding conditions presents a problem in the context of the fluorescence quenching experiment since it means that the ionic quencher is not a nonperturbing probe. To overcome this difficulty, an alternative method was used, in which refolding is first initiated by dilution of the denatured enzyme with the usual acetate refolding buffer and the chloride or iodide ions are then added in a second dilution, after different refolding times. The fluorescence intensity is then measured within 2 ms of the second mix. By varying the refolding delay, this technique allows the quenching profile during folding to be reconstructed while avoiding the effect of the quenching agent on the refolding kinetics (Figure 3c).

The extent of fluorescence quenching is conveniently described by the parameter I^0/I, which is the ratio of the fluorescence intensity in the presence of 0.3 M chloride to that observed in the presence of 0.3 M iodide. At the earliest measurable refolding time (5.1 ms), I^0/I was around 1.8. The effectiveness of a quencher is dependent on its mobility in the solution, and I^0/I therefore is expected to vary with guanidine hydrochloride concentration. It was therefore necessary to predict I^0/I for the fully unfolded state of the protein, under the refolding conditions, by determining it in 6 M guanidine hydrochloride solution and assuming a dependence similar to that for the model peptide Gly-Gly-Trp-Ala which was measured over the full range of denaturant concentration. In this way a value of I^0/I of 2.0 was estimated for the fully unfolded state under the refolding concentrations. This is significantly greater than the value of 1.8 observed at the first measurable time after refolding, which suggests, in accord with the fluorescence intensity data described above, that a significant change in the environment of at least some tryptophan residues occurs within these first few milliseconds of folding. This is a relatively large effect accounting for around a quarter of the total change in quenching observed on refolding.

A second kinetic phase, involving a further and more substantial decrease in fluorescence quenching, is then observed. The time constant of this phase is around 40 ms, comparable to that of the second phase observed in the fluorescence intensity experiment described above. Thus, by about 150 ms the extent of quenching is approaching its limiting value and, within experimental error, no further changes ensue as folding proceeds to completion. Thus, although there is a major change in intrinsic fluorescence intensity with a time constant around 340 ms, this appears not to be accompanied by a significant change in the average accessibility of tryptophan residues as estimated by this solute quenching experiment. To investigate these kinetics more fully, the effect of varying the concentration of the quenching agent was measured at selected times during refolding.

Figure 5 shows Stern–Volmer plots describing the dependence on iodide concentration of the fractional quenching of lysozyme fluorescence at different times during refolding. For a single accessible fluorophore, a simple collisional model predicts a linear relationship between the quencher concentration, [Q], and the change in fluorescence intensity, given by

$$I^0/I = 1 + K_{SV}[Q]$$

where I and I^0 are the intensities in the presence and absence of quencher, respectively, and K_{SV} is the effective quenching constant, which is a ratio of the rate of collisional quenching to the unquenched fluorescence decay rate (Lehrer, 1971; Eftink & Ghiron, 1981). Both of the latter factors could vary during folding, so that the observed variation in iodide

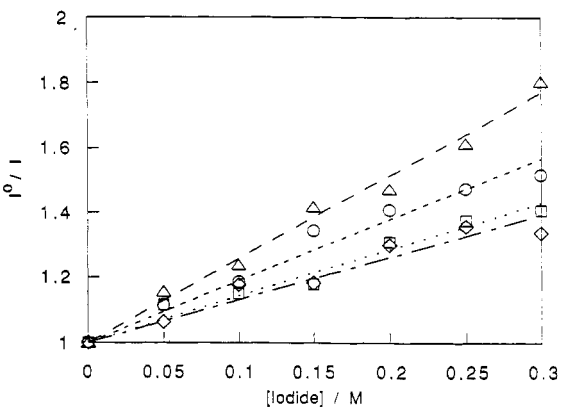

FIGURE 5: Stern–Volmer plots showing the effects of iodide on the intrinsic fluorescence intensity of lysozyme at various times during the refolding reaction. The times were 5 ms (△), 19 ms (○), 100 ms (□), and 2 s (◇). Curvature of the plots, which might be expected for proteins containing more than one tryptophan residue as a result of heterogeneity of the fluorescence, cannot be discerned within the limits of experimental error. The data are therefore fitted to straight lines for simplicity.

quenching during refolding could reflect fluorophore dynamics as well as accessibility. In either case, however, this is clearly a reflection of the development of side-chain interactions in the refolding molecules.

It is clear that by 100 ms the Stern–Volmer plot is already similar to that obtained for the fully refolded protein. However, the MeU-diNAG binding data suggest that only around 30% of molecules are actually in the native state at this time, and even though the intermediates populated by the remaining molecules are less fluorescent, as described above, around half of the total emission (in the absence of quencher) can be attributed to these species. We must therefore conclude that the major changes in solvent accessibility of the individual contributing tryptophans are complete with the formation of these partially organized intermediates. To account for this result, we must conclude that the slow change observed in total fluorescence intensity described above results from increases in quantum yield affecting fluorophores of varying accessibility more or less equally.

ANS Binding. The enhanced fluorescence of ANS when it binds to exposed hydrophobic regions of partially folded protein molecules has been used extensively to detect intermediates in protein folding (Ptitsyn *et al.*, 1990; Semisotnov *et al.*, 1991). The time dependence of fluorescence of ANS during the folding of lysozyme was measured both by inclusion of the dye in the refolding buffer and by its addition in a double-mix experiment after folding had reached various stages of completion. The curves obtained (Figure 6) are qualitatively similar. First, during the dead time of the experiment, a large increase in fluorescence enhancement is observed relative to that in the presence of a fully unfolded state (e.g., in 6 M guanidinium chloride or when the disulfide bridges are cleaved). Interestingly, there is a substantially greater fluorescence enhancement when the ANS is added in the double-mix experiment than when the dye is included directly in the refolding buffer. Thus, implementation of the double-mix technique is again very important since ANS clearly perturbs the refolding kinetics of lysozyme, presumably because its binding to partially folded forms affects their distribution and stabilities.

In the single-mix experiment the fluorescence intensity initially rises in the first few milliseconds and then decays much more slowly. A qualitatively similar initial increase in intensity was also observed in the double-mix experiment at all refolding times, including when the dye was mixed with

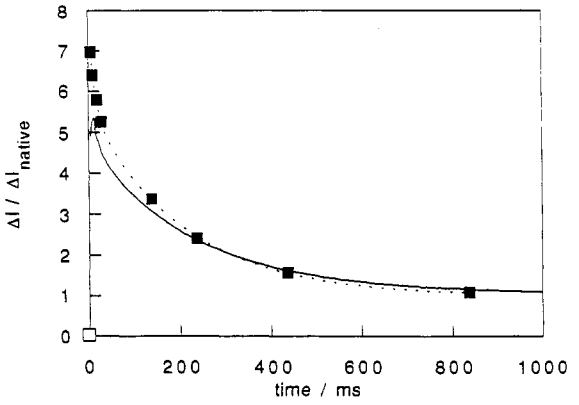

FIGURE 6: Time course of refolding of lysozyme in 0.54 M Gdn-HCl, pH 5.2, 20 °C, monitored by binding to ANS. ΔI is the difference in fluorescence intensity observed between solutions of ANS with and without lysozyme. The parameter plotted, $\Delta I/\Delta I_{native}$, is therefore the ratio of the fluorescence enhancement in the presence of the refolding protein to that observed when the protein is restored to its native state. The continuous line was the trace obtained in an experiment in which the dye was included in the refolding buffer, whereas the individual points (■) were obtained by double-mix experiments in which the dye was added after various periods of refolding (see text). The data at short times (t < 20 ms) are complicated by the fact that binding of the dye to the protein occurs on a similar time scale under these conditions. These data have not been included in fitting the results. The dashed line is a fit of a sum of two exponentials to the data from the double-mix experiment. Also shown (□) is the nominal value of the enhancement at the initiation of refolding: lysozyme denatured in 6 M guanidine hydrochloride causes no ANS fluorescence enhancement, and we have assumed that the same would be true for the fully unfolded state under the refolding conditions.

the native enzyme. This presumably reflects the kinetics of binding of the dye to the protein, rather than any conformational changes in the protein itself. This suggests that the kinetics of binding of ANS may limit the observed fluorescence enhancement and that the species formed very early in folding could in principle, therefore, exhibit a substantially larger enhancement than that measured in this experiment. This initial binding phase was, therefore, omitted from the kinetic analyses, and the subsequent phases were fitted to a sum of two exponentials. In the single-mix experiment time constants of 26 and 260 ms were obtained, the slower phase having an approximately 2.5-fold larger amplitude. The data obtained by the double-mix method are qualitatively similar, but the time constants of both phases are somewhat shorter (14 and 210 ms, respectively), suggesting that binding of the dye to early partially folded intermediates could retard their subsequent transformation into more native-like species. While the observation of a measurable fast phase suggests that there are changes either in binding affinity or in quantum yield in forming later intermediates, the fact that the major intensity decrease occurs only as the slowest step suggests that the majority of intermediates retain significant affinity for ANS and, therefore, still have relatively exposed hydrophobic cores. It is clear, however, that the species formed early in the pathway bind, and enhance the fluorescence of ANS, to a much greater extent than species formed later during folding.

DISCUSSION

Table 1 summarizes the time constants and amplitudes of the kinetic folding phases observed by the four different fluorescent probes and by near-UV CD, as described here, together with data obtained previously from far-UV CD and hydrogen-exchange protection studies (Radford *et al.*, 1992). Comparison of the data from the different experiments allows

Table 1: Kinetic Parameters for Lysozyme Refolding at 20 °C, pH 5.2

experiment	parameter	time constant (ms)	amplitude[a] (%)
exchange protection	α-domain	5 ± 3	48 ± 14
	α-domain	65 ± 25	52 ± 30
	β-domain	11 ± 5	29 ± 8
	β-domain	340 ± 180	71 ± 25
far-UV CD	225 nm	<1	80 ± 10
		25 ± 12	10 ± 10
		345 ± 70	10 ± 10
near-UV CD	289 nm	330 ± 60	100
intrinsic fluorescence intensity		<1	+120 ± 10
		25 ± 3	+290 ± 5
		340 ± 20	−310 ± 5
fluorescence quenching		<1	40 ± 10
		40 ± 6	60 ± 10
ANS binding		<2	+690
		23 ± 5	−120
		250 ± 60	−470
MeU-diNAG binding		350 ± 50	100

[a] In all of these experiments the amplitudes quoted are normalized to the total change between the initiation of refolding and 2 s, by which time all phases discussed in this paper are essentially complete. Thus, small-amplitude slow phases ($t \gg 1$ s), probably attributable to proline isomerism, are not included. Changes within the dead time (t < 1 ms) are calculated from the expected value of the parameter for a fully unfolded state under the refolding conditions (see text).

us to explore the properties of the refolding molecules at various stages along the folding pathway.

Dead Time Folding Events (<5 ms). In the earliest measurable stages of folding a substantial proportion of secondary structure is formed. Reconstruction of the CD spectrum, albeit under refolding conditions slightly different from those employed here, has suggested that the average complement of secondary structure may in fact be very close to that of the native state at this time (Chaffotte *et al.*, 1992). The absence of significant near-UV CD at this stage suggests that few fixed tertiary interactions have yet developed, but various fluorescence parameters suggest, nonetheless, that substantial hydrophobic collapse has already occurred. Thus, the diminished fluorescence intensity suggests that intramolecular quenching mechanisms have already become significant for some, at least, of the tryptophan residues, presumably as a consequence of interactions between side chains in a partially collapsed state. The existence of loose hydrophobic clusters could also explain both the reduction in quenching by iodide ions and the maximal enhancement of ANS fluorescence which are observed during the dead time of the experiment. These species formed early in folding do not exhibit significantly retarded amide exchange, suggesting that they are relatively unstable and rapidly fluctuating between a variety of different conformational states, some at least of which offer little protection from exchange. These features are qualitatively similar to those of equilibrium molten globule states, such as that of the homologous protein α-lactalbumin (Dolgikh *et al.*, 1985; Kuwajima, 1989; Baum *et al.*, 1989; Chyan *et al.*, 1993). Similarly very rapid formation of secondary structure has now been observed in the refolding of a wide range of proteins (Ptitsyn *et al.*, 1990; Kuwajima *et al.*, 1991; Elöve *et al.*, 1992; Mann & Matthews, 1993), but it is harder to generalize about the extent to which this is accompanied by hydrophobic clustering. Intrinsic fluorescence measurements have now provided evidence for early formation of side-chain interactions in lysozyme and several other proteins (Garvey *et al.*, 1989; Khorasanizadeh *et al.*, 1993). On the other hand, ANS

binding seems to develop more slowly in many cases than it does during the refolding of lysozyme (Ptitsyn *et al.*, 1990).

Intermediate Phases (τ < 100 ms). Following these very early folding events, a number of further changes are resolved that are clearly faster than formation of the native structure and therefore must reflect the population of additional folding intermediates. Time constants assigned to these phases by fitting the data to multiple exponential functions vary over a substantial range, from 5 to 65 ms. This variation may in part be a consequence of the limitations of fitting these complex kinetic data to relatively simple functions, but they may also reflect the fact that a range of distinct events is taking place on this time scale and that the various techniques respond differently to different molecular events. The hydrogen-exchange labeling studies, for example, reveal at least two phases of structure formation during this time, giving protection of both α- and β-domains in about 20% of molecules with a time constant around 10 ms and of the α-domain alone in the remaining molecules in two phases with time constants around 10 and 60 ms (Radford *et al.*, 1992; Miranker *et al.*, 1993). The fluorescence and CD measurements represent averages over the different chromophores within each molecule and over the population of molecules in solution, and the kinetic analyses applied to them can only give an indication of the time scale of individual processes. The fact that phases with time constants intermediate between those seen in the pulse labeling experiments are observed by the different techniques for this phase of folding (Table 1) is consistent with this view. Nonetheless, there are clear spectroscopic changes within the period when hydrogen-exchange protection develops, and we can put the data together to describe the average properties of the refolding molecules at various times during the folding process.

Extensive protection of amide hydrogens takes place during these first 100 ms or so of refolding. Within 20 ms protection of amides in the α-domain alone is observed in about 30% of molecules and about 20% of molecules are protected in both the α- and β-domains. Interestingly, hydrogen-exchange labeling in combination with electrospray mass spectrometry (Miranker *et al.*, 1993) suggests that for at least the majority of molecules assembly of the β-domain does not occur independently of, but only in concert with, the α-domain in this folding step. The present results demonstrate that despite this coexistence of both domains in the same molecule on this time scale, this species is not yet fully refolded, as it does not have the capability of inhibitor binding or the near-UV CD characteristic of the native state. There are, therefore, at least two distinct types of intermediate formed in this phase, one protecting amides throughout the molecule and the other protecting amides only in the α-domain. These are not formed consecutively but occur, rather, on alternative folding pathways. Further, the intermediates in which the α-domain only is protected are formed via two distinct pathways, one with a $\sim$5 ms and the other with a $\sim$60 ms time constant. By 100 ms, around 30% of molecules have reached the native state and virtually all of the remainder populate one or the other of these intermediates.

The data described in this paper show that the progressive protection of backbone amides from exchange is paralleled by the development of side-chain interactions which presumably stabilize the backbone folding. This point is also illustrated by results obtained for other proteins (Elöve *et al.*, 1992; Matouschek *et al.*, 1992). The species in which protection first becomes significant seem to have tighter clustering of hydrophobic side chains than was present in the very early intermediates, as reflected by a diminution in ANS fluorescence enhancement, the development of exchange protection of at least one tryptophan indole hydrogen, and the observation that the quenching of fluorescence by iodide ions is approaching its limiting value by 100 ms, suggesting that those tryptophans resistant to quenching in the native state are already resistant in the intermediate species populated at this stage.

Despite these features, these intermediates are clearly still not fully native-like, even within the highly stabilized α-domain. In particular, the intrinsic fluorescence intensity is much smaller than in either the native or the fully denatured state and the near-UV CD spectrum has not developed by this stage. These observations are consistent with tryptophan side-chain interactions that are extensive but not yet specific, providing ample opportunities for interresidue energy transfer which thereby diminishes fluorescence. Overall, therefore, we would conclude that in the intermediates prevailing at around 100 ms the α-domain still has some characteristics of the molten globule model, notably in its relatively nonspecific side-chain interactions, although its hydrophobic core is much more highly developed than in earlier intermediates. This is supported by the fact that these intermediates bind ANS to a much lesser extent than earlier intermediates. Similarly, this intermediate is much more structured than the A-state of α-lactalbumin, for example, where there is enhanced accessibility to iodide fluorescence quenching (Lala & Kaul, 1992) and much more labile secondary structure as indicated by the much lower exchange protection factors (Chyan *et al.*, 1993).

Slow Phases (τ < 250 ms). The native conformation of the enzyme, as detected by binding of a competitive inhibitor, forms with a single time constant of around 350 ms. A very similar time development is observed for the near-UV CD spectrum which reflects the fixing of aromatic chromophores in an asymmetric environment. Two of these residues (Trp 62, 63) form part of the active site cleft in native lysozyme, and to this extent the similarity with the inhibitor binding result is not surprising. The majority, however, form the core of the α-domain and the probable implication of this is that the tertiary structure of this domain remains labile until it is organized cooperatively with the β-domain into the final native structure. Thus, in spite of the heterogeneity of the folding mechanism, the rates at which molecules finally acquire their native structures appear to be remarkably uniform. This is a dramatic result in view of the fact that around 20% of molecules have formed stable structures in both the α- and β-domains within the first 20 ms of folding. Clearly, the coexistence of both folding domains in a single molecule is not sufficient to ensure facile organization into the final native structure even though the individual domains presumably have adopted significantly native-like structures. Rather, there must be an additional step which involves their integration into a cooperative structure, and this must effectively be rate limiting. A similar situation, in which the late stages of folding involve the organization of interactions between substructures that are already well developed, has also been observed in other proteins (Tweedy *et al.*, 1990; Mann & Matthews, 1993; Ballery *et al.*, 1993). The apparent kinetic homogeneity of native structure formation in hen lysozyme suggests, therefore, that all molecules may have to surmount a common rate-limiting organizational barrier, even though they have differed in earlier stages of their refolding mechanisms.

The rate of native structure formation is close to that of the major phase of protection of amides in the β-domain, suggesting that in these molecules organization of this domain occurs effectively concurrently with its integration into the native structure. Various other spectroscopic parameters also recover

to their limiting native values with rates very similar to this, offering further insight into the nature of the structural changes involved at this stage of refolding. In general, these parallel changes support the idea that this slow step involves reorganization and fixing of side-chain interactions, probably in a rather cooperative fashion. It is evident that this occurs from intermediates which are themselves persistently structured, implying that the origin of this slow step may be the barriers to reorganization generated in highly structured states.

Although the nature of this final step remains unclear, some speculation is possible on the basis of the data presented here. The retardation of this phase, but not of the fast phase detected by intrinsic fluorescence, with increasing ionic strength suggests that formation of specific electrostatic interactions between charged side chains occurs in this late phase of folding. The final loss of ANS fluorescence enhancement also occurs in this phase, presumably reflecting tight interlocking of hydrophobic side chains, and the recovery of the intrinsic fluorescence intensity may also reflect rearrangement of tertiary interactions as the native state forms. Another possible cause of the latter would be desolvation of the indole fluorophores, although the observations that both susceptibility to iodide quenching and resistance to indole NH exchange are already native-like in earlier intermediates may suggest that the distribution of at least the buried tryptophan residues in relation to the bulk solvent does not change significantly in this slow phase.

It seems likely that in a significant proportion of molecules, at least, this slow phase also involves breakdown and reorganization of some nonnative interactions. Thus, additional negative ellipticity develops in the early stages of folding and tryptophan fluorescence becomes more highly quenched than in either the native or unfolded states. The additional ellipticity has been suggested to arise from some form of nonnative structure, perhaps involving a disulfide bridge in an incorrect configuration (Chaffotte *et al.*, 1992). The return of the ellipticity and fluorescence intensity to their native-state values occurs on the time scale characteristic of the slowest step in folding.

Taken together with the results of the refolding studies of lysozyme by hydrogen-exchange labeling and other techniques, the fluorescence studies reported here add a new dimension to our understanding of the folding pathway. There is now firm evidence for at least three distinct folding phases. First, there is an initial collapse to a state possessing extensive secondary structure but only weak tertiary interactions. This is followed by a reorganization into at least two different intermediates which now show substantial protection from hydrogen exchange; one of these has persistent structure in both domains while in the other, major, species only the α-domain is sufficiently stable to afford significant protection. In neither of these species are the protein side chains yet fully organized into the specific configurations characteristic of the native functional enzyme, however. Only in the final slowest phase of folding does this occur; this step involves final rearrangement of at least some of the residues in the hydrophobic core, development of specific ionic interactions, and the docking of the two folding domains to form the active site cleft.

ACKNOWLEDGMENT

We thank Simon Lever and Michael Carey of Applied Photophysics Ltd. for access to and assistance with their stopped-flow system.

REFERENCES

Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84–91.

Baldwin, R. L., & Roder, H. (1991) *Curr. Biol. 1*, 218–220.

Ballery, N., Desmadril, M., Minard, P., & Von, J. M. (1993) *Biochemistry 32*, 708–714.

Blake, C. C. F., Koenig, D. F., Mair, G. A., & Sarma, R. (1965) *Nature 206*, 757–761.

Chaffotte, A. F., Guillou, Y., & Goldberg, M. E. (1992) *Biochemistry 31*, 9694–9703.

Chyan, C. L., Wormald, C., Dobson, C. M., Evans, P. A., & Baum, J. (1993) *Biochemistry 32*, 5681–5691.

Creighton, T. E., Ed. (1992) *Protein Folding*, Freeman, New York.

Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Gilmanshin, R. L., Lebedev, O. Y., Semisotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109–121.

Eftink, M. R. (1991) *Methods Biochem. Anal. 35*, 127–207.

Eftink, M. R., & Ghiron, C. A. (1981) *Anal. Biochem. 114*, 199–227.

Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. E. (1992) *Biochemistry 31*, 6876–6886.

Englander, S. W., & Mayne, L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 243–265.

Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol. 3*, 75–83.

Garvey, E. P., Swank, K., & Matthews, C. R. (1989) *Proteins: Struct., Funct. Genet. 6*, 259–266.

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., & Rupley, J. A. (1972) in *The Enzymes* (3rd ed.) (Boyer, P. D., Ed.) Vol. 7, pp 665–868, Academic, Orlando, FL.

Jennings, P. A., Saalau-Bethell, S. M., Finn, B. E., Chen, X., & Matthews, C. R. (1991) *Methods Enzymol. 202*, 113–126.

Kato, S., Okamura, M., Shimamoto, N., & Utiyama, H. (1981) *Biochemistry 20*, 1080–1085.

Khorasanizadeh, S., Peters, I. D., Butt, T. R., & Roder, H. (1993) *Biochemistry 32*, 7054–7063.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Kuwajima, K. (1989) *Proteins: Struct., Funct. Genet. 6*, 87–103.

Lala, A. K., & Kaul, P. (1992) *J. Biol. Chem. 267*, 19914–19918.

Lehrer, S. S. (1971) *Biochemistry 10*, 3254–3263.

Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32*, 5282–5290.

Matthews, C. K. (1991) *Curr. Opin. Struct. Biol. 1*, 28–35.

Matouschek, A., Serrano, L., Meiering, E. M., Bycroft, M., & Fersht, A. R. (1982) *J. Mol. Biol. 224*, 837–845.

Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., & Dobson, C. M. (1993) *Science 262*, 896–899.

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

Radford, S. E., Dobson, C. M., & Evans, P. A. (1992) *Nature 358*, 302–307.

Segawa, S., & Sugihara, M. (1984) *Biopolymers 23*, 2489–2498.

Semisotnov, G. V., Rodionova, O. I., Razgulaev, Uversky, V. N., Gripas, A. F., & Gilmanshin, R. L. (1991) *Biopolymers 31*, 119–128.

Smith, C. J., Clarke, A. R., Chia, W. N., Irons, L. I., Atkinson, T., & Holbrook, J. J. (1991) *Biochemistry 30*, 1028–1036.

Tweedy, N. B., Hurle, M. R., Chrunyk, B. A., & Matthews, C. R. (1990) *Biochemistry 29*, 1539–1545.

Van Landshoot, A., Lootiens, F. G., Clegg, R. M., Sharon, N., & De Bruyne, C. K. (1977) *Eur. J. Biochem. 79*, 275–283.

Yang, Y., & Hamaguchi, K. (1980) *J. Biochem. 87*, 1003–1014.