Mechanistic Studies of the Folding of Human Lysozyme and the Origin of Amyloidogenic Behavior in Its Disease-Related Variants†

Denis Canet,‡ Margaret Sunde,‡ Alexander M. Last,‡ Andrew Miranker,‡,§ Andrew Spencer,‖ Carol V. Robinson,‡ and Christopher M. Dobson*,‡

Oxford Centre for Molecular Sciences, University of Oxford, New Chemistry Laboratory, South Parks Road, Oxford OX1 3QT, U.K., and Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, U.K.

Received December 23, 1998; Revised Manuscript Received February 18, 1999

ABSTRACT: The unfolding and refolding properties of human lysozyme and two amyloidogenic variants (Ile56Thr and Asp67His) have been studied by stopped-flow fluorescence and hydrogen exchange pulse labeling coupled with mass spectrometry. The unfolding of each protein in 5.4 M guanidine hydrochloride (GuHCl) is well described as a two-state process, but the rates of unfolding of the Ile56Thr variant and the Asp67His variant in 5.4 M GuHCl are ca. 30 and 160 times greater, respectively, than that of the wild type. The refolding of all three proteins in 0.54 M GuHCl at pH 5.0 proceeds through persistent intermediates, revealed by multistep kinetics in fluorescence experiments and by the detection of well-defined populations in quenched-flow hydrogen exchange experiments. These findings are consistent with a predominant mechanism for refolding of human lysozyme in which one of the structural domains (the α-domain) is formed in two distinct steps and is followed by the folding of the other domain (the β-domain) prior to the assembly of the two domains to form the native structure. The refolding kinetics of the Asp67His variant are closely similar to those of the wild-type protein, consistent with the location of this mutation in an outer loop of the β-domain which gains native structure only toward the end of the refolding process. By contrast, the Ile56Thr mutation is located at the base of the β-domain and is involved in the domain interface. The refolding of the α-domain is unaffected by this substitution, but the latter has the effect of dramatically slowing the folding of the β-domain and the final assembly of the native structure. These studies suggest that the amyloidogenic nature of the lysozyme variants arises from a decrease in the stability of the native fold relative to partially folded intermediates. The origin of this instability is different in the two variants, being caused in one case primarily by a reduction in the folding rate and in the other by an increase in the unfolding rate. In both cases this results in a low population of soluble partially folded species that can aggregate in a slow and controlled manner to form amyloid fibrils.

The in vitro refolding of several *c*-type lysozymes has been studied in great detail, and many of the key steps in the folding reactions have been defined (*1, 2*). Hen and human lysozymes in particular have been investigated by a wide range of complementary biophysical techniques, each designed to probe different features of the recovery of the nativelike properties. These techniques have involved a variety of stopped-flow methods, including both optical (*3*−*8*) and NMR procedures (*9*), and hydrogen exchange pulse labeling experiments monitored by NMR and mass spectrometry (*3, 8, 10, 11*). The results indicate that hen and human lysozymes refold in the presence of their native disulfide bonds via multiple pathways and through a series of well-defined intermediates. In both proteins a key feature appears to be the formation of persistent nativelike structure in each of the two structural domains (α and β) prior to docking of the domains to generate the native close-packed structure (*12, 13*).

In most molecules the α-domain forms persistent structure during folding more rapidly than the β-domain. In the case of the hen protein the persistent overall fold of the α-domain is formed in a highly cooperative manner prior to the folding of the β-domain (*3*). In the case of human lysozyme, however, two well-defined steps are seen (*8*). In the first step, two of the four α-helices fold very rapidly, followed by the other two on a slower time scale. The rate of this latter step is comparable to that of the formation of the β-domain. In addition to the major folding events, slower processes have been detected for both proteins in fluorescence experiments. Although some evidence suggests that these can involve proline isomerization (*14*), mutation experiments with the human protein show that other factors may also be important. For example, disulfide bond isomer-

† D.C. was supported by the Royal Society and the European Community; M.S. was supported by the Medical Research Council and Lady Margaret Hall, Oxford; C.V.R. was supported by a Royal Society Research Fellowship. This work is a contribution from the Oxford Centre for Molecular Sciences, which is supported by EPSCR, BBSRC, and MRC. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Institute and the Wellcome Trust.

* Corresponding author. Tel: +44-(0)1865 275916. Fax: +44-(0)1865 275921. E-mail: chris.dobson@chem.ox.ac.uk.

‡ University of Oxford.

§ Present address: Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Ave., P.O. Box 208114, New Haven, CT 06520-8114.

‖ Institute of Food Research.

10.1021/bi983037t CCC: $18.00
Published on Web 04/30/1999

ization has been suggested to be responsible for the slowest detectable steps in the refolding of human lysozyme (*15*).

In recent work, two single-point mutations of human lysozyme have been identified as the origin of hereditary systemic amyloidosis (*16*), one of the family of amyloid diseases that includes Alzheimer's disease and the spongiform encephalopathies including Creutzfeldt-Jakob disease. The variant lysozymes are enzymatically active but are able to form fibrils in vitro as well as in vivo (*17*). The amyloidogenic nature of the human lysozyme variants has been associated with perturbations to the stability and dynamics of the proteins resulting from the mutations (*17*). Hydrogen exchange measurements probed by mass spectrometry have shown that both variant proteins undergo transient unfolding (*17*). Denaturation studies show that the variant proteins have a decreased thermal stability relative to the wild-type protein and that, unlike the wild-type protein, partially structured states are significantly populated under equilibrium conditions. It has been suggested that such species are associated with the aggregation process which results in amyloid fibrils (*17*). In addition, stopped-flow fluorescence studies of the folding and unfolding of the Ile56Thr variant show that a kinetic intermediate is more highly populated in the variant protein than in the wild-type protein (*18*).

In this paper we explore the folding and unfolding properties of both amyloidogenic human lysozyme variants and compare them in detail with those of the wild-type protein. The perturbations induced by the mutations shed light on the nature of the structural transformations taking place during the folding of human lysozyme with its disulfide bonds intact. The folding of lysozyme in vivo occurs in the endoplasmic reticulum and is associated with the formation of the four disulfide bonds characteristic of the native state (*19*). In this regard, the in vitro folding studies of the oxidized protein do not correspond directly to the *de novo* folding occurring in a cell. With regard to amyloid formation, however, studies have shown that ex vivo Asp67His fibrils can be solubilized and that the intact, monomeric Asp67His lysozyme recovered by such a procedure is enzymatically active (*17*). This indicates that the variant protein reaches its native form with the correct disulfide bonds upon de novo folding in the cell, and that subsequent deposition of the fibrillar form in tissues such as the liver involves a disulfide-intact species. The in vitro folding studies of the oxidized proteins, therefore, are those appropriate to probe the transitions that result in the formation of fibrils in the lysozyme amyloidoses. In addition, the insight from such a study is likely to be relevant to similar processes occurring in the other proteins associated with this family of related diseases (*20*).

MATERIALS AND METHODS

Production and Purification of Recombinant and Variant Lysozymes. The coding sequence for human lysozyme (*21*) and the cDNAs of the Ile56Thr variant and the Asp67His variant (*17*) were cloned into the GAM-fusion vector pIGF (*22*) and transferred into *Aspergillus niger* AB4.1 by co-transformation (*23*). Selection for transformed protoplasts was on *Aspergillus* minimal medium with 1.2 M sorbitol at 25 °C. Transformed colonies were then assayed for lysozyme secretion by plate assays. Individual lysozyme-producing colonies were grown in 100 mL of polyvinylpyrrolidone (PVP)[1] complete medium (*24*) and assayed for lysozyme activity (*25*). The highest yielding individual was used as the strain for lysozyme production in white soya milk medium. The lysozyme was harvested from the culture medium by ion exchange chromatography and was further purified by gel filtration.

Expression levels of variant proteins, although substantial, were less than that of the wild-type protein, which was probably due to their higher susceptibility to proteolysis. Producing the variant lysozymes at 17.5 °C after an initial starter period of 2 days at 25 °C improved the yield of protein. These yields were up to 57 mg/L for the wild-type protein, 30 mg/L for the Asp67His variant, and 18 mg/L for the Ile56Thr variant. Mass spectrometry was used to check the integrity of the recombinant proteins: the measured masses of the Asp67His and the Ile56Thr variant were 14 714 $\pm$ 1 and 14 680 $\pm$ 2 Da respectively, consistent with their calculated masses from the amino acid sequence (14 715 and 14 680 Da, respectively). Human lysozyme uniformly labeled in ^{15}N was produced in PVP labeling medium at 20 °C (*26*). This is a minimal medium where the sole source of nitrogen is $^{15}NH_4Cl$ (Cambridge Isotope Laboratories). Purification of the labeled protein from the culture medium was by ion exchange chromatography. The measured mass of the ^{15}N-labeled protein was 14 891 $\pm$ 2 Da, confirming greater than 99% incorporation of ^{15}N (14 893 Da predicted for 100% labeling).

Stopped-Flow Fluorescence Experiments. Unfolding and refolding experiments were carried out using a Bio-Logic (Claix, France) SFM3 mixer with a cell path length of 1.5 mm and a Bio-Logic PMS 200 detection system. The theoretical dead time based on the flow rate was 2 ms. The excitation wavelength was set at 280 nm, with slit widths at 4 nm. The emission was filtered with a WG 320 nm cutoff filter. Data acquisition was performed using two time bases in order to record a higher number of points in the initial times following mixing. This results in the traces shown on the figures having a higher noise level (the signal being less filtered) at the beginning of the kinetic profiles. For unfolding experiments, 1 volume of protein in buffer solution (20 mM acetate at pH 5.0) was mixed with 10 volumes of a solution containing 6 M GuHCl at pH 5.0. For refolding experiments, 1 volume of protein in 6 M GuHCl was mixed with 10 volumes of buffer solution (20 mM acetate at pH 5.0). The final protein concentration in the cell was 1 μM. Exact concentrations of GuHCl were measured from the refractive index of the solutions (*27*).

The kinetic traces were analyzed as sums of exponentials with Bio-Kine software supplied by Bio-Logic, as well as with Origin software from Microcal Software, Inc. (Northampton, MA). For the unfolding experiments eight shots were fitted individually. The numerical values given are the means and standard deviations of these eight fits. The traces displayed are the averages of the eight shots. For the refolding experiment of the Asp67His variant, 27 shots were

[1] Abbreviations: GuHCl, guanidine hydrochloride; GuDCl, deuterated guanidine hydrochloride; PVP, polyvinylpyrrolidone; pH*, pH of deuterated solution uncorrected for isotopic effect (meter reading); rmsd, root-mean-square deviation.

Table 1: Kinetic Parameters for Refolding of Human Lysozyme and Two Amyloidogenic Variants Followed by Stopped-Flow Fluorescence and Hydrogen Exchange Pulse Labeling Coupled with Mass Spectrometry

		stopped-flow fluorescence		quenched-flow/mass spectrometry
		amplitude (%)	time constant (ms)	time constant (ms)
fast (phase I)	Ile56Thr	28 ± 3	7.5 ± 0.1	<3.5 for I_1
	Asp67His	41	5.5	and ca. 6.0 for
	wild type	43 ± 7	7.3 ± 0.8	the fast track
intermediate (phase II)	Ile56Thr	59 ± 8	399 ± 37	502 ± 28
	Asp67His	50	50	51 ± 6
	wild type	52 ± 6	50 ± 10	56 ±7
slow (phase III)	Ile56Thr	12 ± 5	1605 ± 584	≤10% of
	Asp67His	9	825	missing
	wild type	5 ±3	345 ± 40	amplitude

averaged and fitted. For the refolding traces of the wild-type protein and the Ile56Thr variant, the experiment was repeated on an Applied Photophysics SX.17MV stopped-flow fluorimeter (Leatherhead, U.K.), and the numerical values given are the means and standard deviations of experiments with different preparations of protein (four in the case of the wild-type protein, two on the Bio-Logic stopped-flow and two on the Applied Photophysics stopped-flow, and two in the case of the Il56Thr variant, one on each machine).

Quenched-Flow Hydrogen Exchange Pulse Labeling. Hydrogen exchange pulse labeling experiments were carried out using a Bio-Logic QFM5 mixer. GuHCl was deuterated by three cycles of dilution in D_2O followed by lyophilization. The two variant proteins and ^{15}N-labeled wild-type protein were mixed in equal proportions and dissolved in a 6 M GuDCl solution. One volume of this solution was mixed with 10 volumes of the refolding buffer solution (20 mM acetate in H_2O, pH 5.0), and refolding was allowed to proceed for various times between 3.5 ms and 4 s at 20 °C. The resulting solutions were then mixed with 5 volumes of buffer solution (200 mM borate in H_2O, pH 10.0) in the pulse labeling step, and exchange was allowed to take place for 8.4 ms. Then the reactions were quenched by mixing the solutions with 5 volumes of an acetic acid solution in H_2O whose concentration was adjusted to between 0.5 and 1 M to result in a pH of 3.8 after the quench. The fully exchanged samples were prepared by mixing 1 volume of protein in a 6 M GuDCl solution with 15 volumes of refolding buffer in H_2O solution, followed by heating to 80 °C for 15 min.

Mass Spectrometry. The samples from the pulse labeling procedure were collected and placed on ice. They were then extensively washed with deionized water at pH 3.8 (adjusted by addition of formic acid) in Centricon-10 concentrators (Amicon) at 4 °C. Mass spectra were recorded on a Micromass Platform II spectrometer. Samples were introduced using a nanoflow electrospray probe with borosilicate glass capillaries prepared in house (*28*). 2 μL of solution containing approximately 20 μM protein was used for each spectrum. All mass spectra represent the +11 charge state with minimal smoothing, expressed on a mass scale. Data analysis was performed on transformed mass spectra using Win-IR (Bio-Rad/Galactic Industries Co., Salem, NH). The spectra were fitted with a linear baseline and a combination of Gaussian distributions to determine the relative proportions of the different components. The peak widths at half-height for all the native species were reproducibly 12.7 ± 0.7 Da, which gives an estimate of the resolution in these experiments. The calculated areas for each time point were fitted simultaneously to extract kinetic parameters using Sigmaplot (SPSS Inc., Chicago, IL). The errors given in Table 1 are the standard errors of these fits.

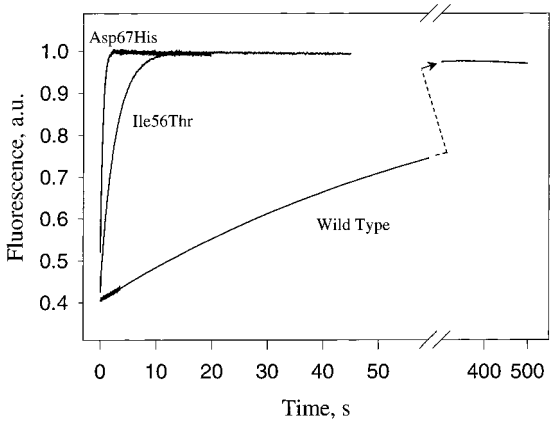


FIGURE 1: Unfolding kinetics followed by stopped-flow fluorescence of recombinant wild-type human lysozyme and the Asp67His and Ile56Thr variants at pH 5.0, 5.4 M GuHCl and 20 °C. The solid lines represent single exponentials of time constants 0.44 ± 0.02, 2.6 ± 0.06, and 69.5 ± 1.8 s for the Asp67His variant, the Ile56Thr variant, and the wild-type protein, respectively. The three traces have been scaled to unit amplitude of the fluorescence of the unfolded species.

RESULTS

Stopped-Flow Optical Measurements. The unfolding of the three proteins in a solution of 5.4 M GuHCl at pH 5.0 and 20 °C was followed by stopped-flow fluorescence spectroscopy. The results are shown in Figure 1. The kinetic traces fit well to single-exponential functions with a time constant of 69.5 s (±1.8 s) for the wild-type protein. The unfolding of the variants was much faster than that of the wild type, time constants of 2.60 s (±0.06 s) being measured for the Ile56Thr variant and 0.44 s (±0.02 s) for the Asp67His variant. The unfolding of the Ile56Thr variant under these conditions is therefore approximately 30-fold faster than that of the wild-type protein. This is consistent with a previous report that the denaturation of the Ile56Thr variant is between 1 and 2 orders of magnitude faster than that of the wild-type protein at pH 4.0 and 10 °C (*18*). The difference in the unfolding rates of the Asp67His variant and the wild-type protein is even more dramatic. The Asp67His variant unfolds some 160 times faster than the wild-type protein.

The refolding of the wild-type and variant lysozymes in a solution containing 0.54 M GuHCl at pH 5.0 and 20 °C was

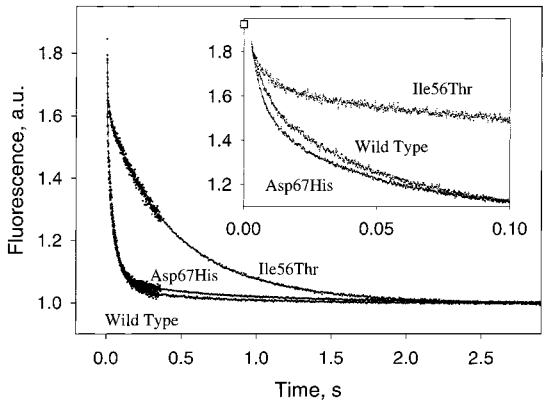


FIGURE 2: Refolding kinetics followed by stopped-flow fluorescence of wild-type human lysozyme and the Asp67His and Ile56Thr variants at pH 5.0, 0.54 M GuHCl and 20 °C. The insert shows the kinetic traces over the first 100 ms, and the white square is the estimated fluorescence signal of the unfolded proteins under the final conditions. The solid lines represent fits to three exponentials with the parameters given in Table 1. The fluorescence levels of the three traces have been adjusted with a multiplying factor so that the native proteins have an arbitrary fluorescence level of 1.0. No offset has been applied.

followed by stopped-flow fluorescence spectroscopy (Figure 2). The data show that all three proteins refold completely under these conditions, and the fluorescence changes from the unfolded state to the native state are identical in each case. The fluorescence intensity of the fully unfolded protein under the refolding conditions was estimated by recording the fluorescence of the fully unfolded state in 6 M GuHCl and reducing this fluorescence by a factor of 1.21. This factor accounts for enhancement of the tryptophan fluorescence by GuHCl and was determined by using the posttransition slope in a separate equilibrium unfolding experiment using apo-α-lactalbumin. The insert in Figure 2 shows that the kinetic traces extrapolate to values consistent with this estimate of the initial fluorescence; i.e., this indicates that no unresolved fluorescence changes take place in the dead time of the experiments. Analysis of the fluorescence traces shows that in each case the kinetic data for the refolding processes fit closely to three exponential steps; details of the fits are given in Table 1. We designate these phases I (fast), II (intermediate), and III (slow). The data (Table 1) show that the refolding rates and amplitudes of all phases for the Asp67His variant are similar to those for the wild-type protein. The amplitudes of the refolding phases of the Ile56Thr variant are similar to those observed in the other two proteins, and the rate of phase I is also similar. This variant, however, displays a much slower rate for phase II and a slightly slower rate for phase III.

Phases II and III can be related to those described previously for wild-type human lysozyme in terms of both the amplitudes and rate constants (*15*). Phase I was not seen previously because it would have occurred in the dead time of the experiments [reported as 10 ms (*15*)]. Thus, for the wild-type protein the most rapid phase observed previously represents 90% of the fluorescence intensity change (*15*) and corresponds to phase II in our experiments. Although this phase represents only 52% of the total fluorescence change in our measurements, it corresponds to 90% of the amplitude change observed in phases II and III together. Moreover, the time constants measured by Herning et al. (*15*) for the

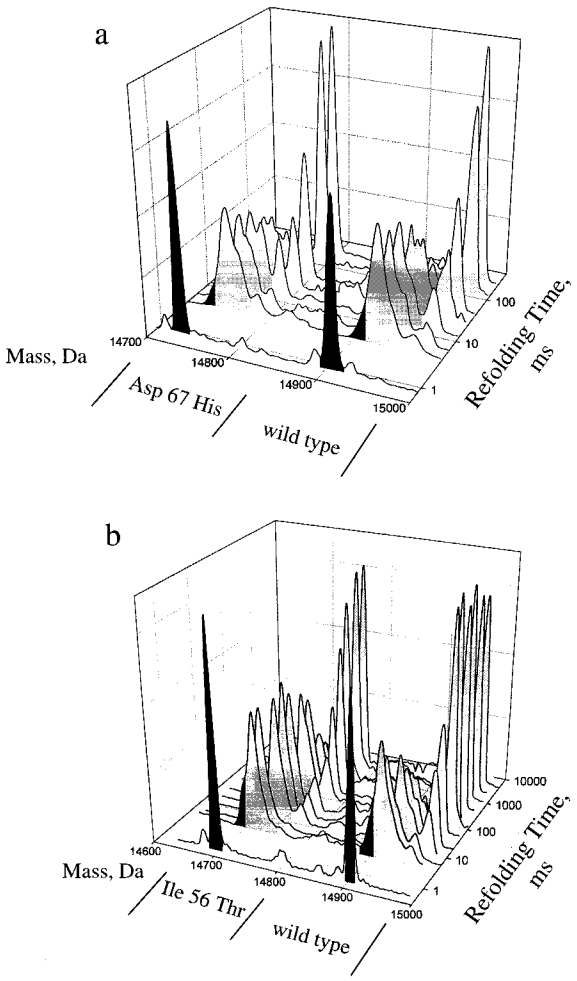


FIGURE 3: Mass spectra showing time evolution of the masses of mixtures of ^{15}N wild-type lysozyme and each of the variant proteins, pulse labeled at different stages of refolding at pH 5.0, 0.54 M GuHCl and 20 °C. The initial spectra (shown before 1 ms) represent protein without incorporation of deuterons (colored black). As the time of refolding increases, additional species are seen which have gained protection against exchange and thus have increased in mass due to incorporation of larger numbers of protected deuterons. These correspond to folding intermediates (colored gray) and nativelike species (colored light gray). (a) Asp67His variant and ^{15}N-labeled wild-type protein. (b) Ile56Thr variant and ^{15}N-labeled wild-type protein.

wild-type protein, at pH 3.0, 25 °C, and 1 M GuHCl, 70 and 625 ms, are consistent with the values of 50 and 345 ms measured in this work for phase II and phase III, respectively, at pH 5.0, 20 °C, and 0.54 M GuHCl. We therefore observe essentially the same refolding kinetics for wild-type human lysozyme as previously described (*15*), augmented by an initial very fast phase (of amplitude 43% and time constant 7.3 ms). In addition, the refolding of the Asp67His variant is similar to that of the wild-type protein while the Ile56Thr variant shows reduced rates for phase II and III.

Quenched-Flow Hydrogen Exchange Pulse Labeling Measurements. The refolding of the variant and the wild-type proteins was studied using hydrogen exchange pulse labeling monitored by electrospray mass spectrometry under the same refolding conditions as for the stopped-flow optical measurements (pH 5.0, 20 °C, and 0.54 M GuHCl) (Figure 3). As the pulse labeling method is highly sensitive to small changes in solution conditions (for example, fluctuations in the pH

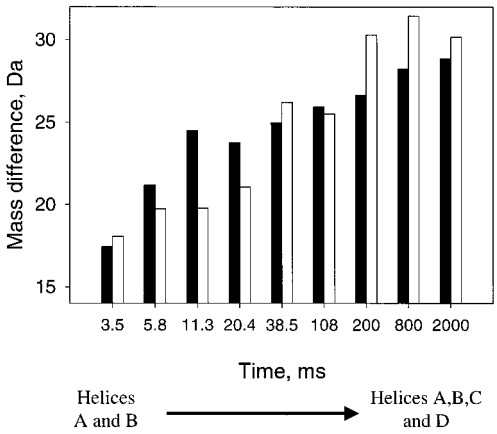


FIGURE 4: Evolution with time of the number of protected hydrogens (relative to the unprotected protein) of the intermediate species in the pulse-labeled refolding experiment from Figure 3b. Data are for the Ile56Thr variant (■) and the wild-type protein (□).

of the labeling pulse), the refolding of each variant protein has been studied simultaneously with that of the wild-type protein using a method described previously (*11*) using wild-type protein uniformly labeled with ^{15}N. This labeling enables the *m*/*z* peaks of the variant and wild-type proteins to be fully separated in the mass spectrometer, enabling their hydrogen exchange behavior to be compared directly. As the time of refolding increases, the proteins gain persistent structure and hence protection against hydrogen exchange. The protected amide hydrogens remain deuterated during the labeling pulse instead of being replaced by protons. This results in the appearance of peaks in the mass spectra having higher masses than that of the fully proton-containing protein.

In experiments with both variants, at the shortest refolding time (3.5 ms), the major peak corresponds to a mass of 17.7 ± 0.3 Da greater than that of the nondeuterated protein. This indicates that persistent structure able to protect this number of amide hydrogens against exchange develops in the dead time of the refolding experiment. As the refolding time increases, this peak shifts progressively to higher mass, and an additional well-resolved peak appears centered at a 64.1 ± 0.5 Da mass higher than that corresponding to the nondeuterated protein; this mass increase corresponds exactly to the number of strongly protected amide hydrogens in the native state (*29*). The variation in intensity of this peak with time is clearly similar for the Asp67His variant and the wild-type protein at all refolding times; see Figure 3a. In Figure 3b, however, the corresponding peak for the Ile56Thr variant gains intensity much more slowly as the refolding time is increased, indicating much slower folding. This correlates with the stopped-flow fluorescence experiments which show that the refolding kinetics of the Asp67His variant and of the wild-type protein are similar but that the refolding of the Ile56Thr variant (phase II) is significantly slower (Figure 2 and Table 1).

The change in mass of the species with intermediate protection was analyzed as a function of time for the Ile56Thr variant in a mixture with the wild-type protein. The average difference in mass between the intermediate and the unprotected species is shown in Figure 4 at different refolding times. For the Ile56Thr variant and the wild-type protein the number of protected amides is 17.7 ± 0.3 at 3.5 ms and increases to 29.7 ± 0.7 as the refolding time increases. These

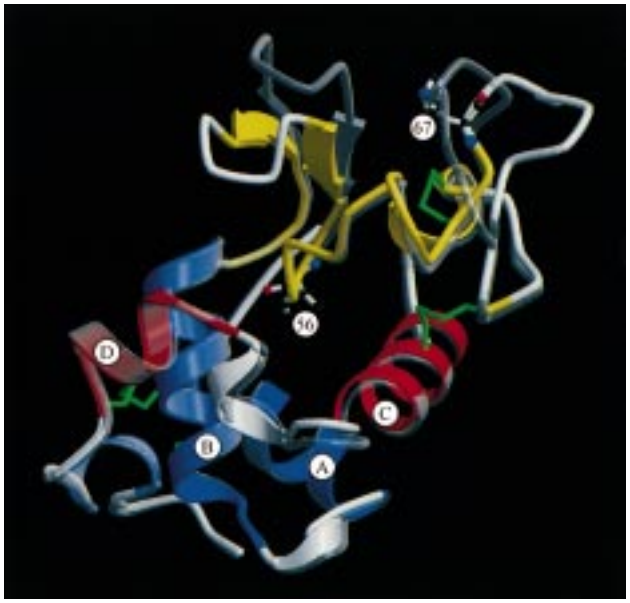


FIGURE 5: Structure of Asp67His shown as a ribbon representation. The chain is colored with the same code as in Scheme 1. The four disulfide bonds are shown in green. The structure of the wild-type protein is superimposed as a semitransparent gray ribbon. The two mutated residues are shown in ball-and-stick representation and the four α-helices are labeled A−D. The figure was generated with the programs Molscript (*34*) and Raster3D (*35*−*37*).

findings are consistent with previous studies (*8*, *11*) in which it was shown by NMR and mass spectrometry that wild-type human lysozyme gains hydrogen exchange protection in α-helices A and B, and the nearby 3_{10} helix, of the α-domain during the dead time of the quenched-flow experiment (3.5 ms). This represents a first folding intermediate (I_1), corresponding to structural elements that protect 18 amides hydrogens in the native state (*29*). A second intermediate (I_2) corresponds to a species with additional protection in helices C and D, i.e., having persistent structure in the complete α-domain (Figure 5). This second intermediate is expected to have 29 protected amide hydrogens according to the hydrogen exchange properties of the native state (*29*). The behavior of the two proteins is essentially identical, as shown in Figure 4, indicating that the folding of the α-domain of the Ile56Thr variant is similar in terms of both the number of protected amides and the rate of protection to that of the wild-type human lysozyme and that the differences noted earlier in the overall folding behavior are associated with the different rates of formation of the *β*-domain in the wild-type protein and Ile56Thr variant.

Figure 6 shows the time course of appearance of the species having nativelike protection in both experiments. This confirms the very close similarity of the formation of the native hydrogen exchange protection in the Asp67His variant and in the wild-type protein (Figure 3a), and the kinetics are consistent with those from stopped-flow fluorescence experiments (Figure 2 and Table 1). Figure 6 also shows the slower development of nativelike hydrogen exchange protection for the Ile56Thr variant than for the wild-type protein. Here again the results are in good agreement with the stopped-flow fluorescence data (Figure 2 and Table 1) which show slower folding of the Ile56Thr variant relative to the wild-type protein. The experimental data in Figure 6 also reveal the presence of ca. 10% of molecules folding

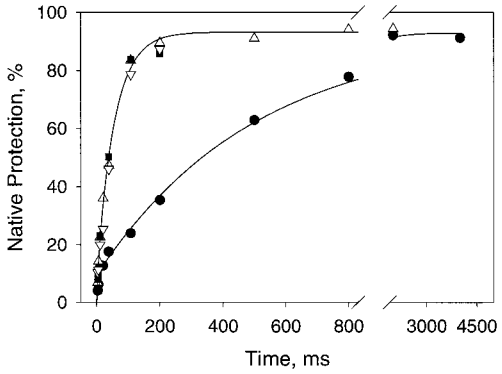


FIGURE 6: Formation of nativelike protection with time in pulse labeling refolding experiments: Asp67His variant (■) and wild-type human lysozyme (▽) from Figure 3a; Ile56Thr variant (●) and wild-type human lysozyme (△) from Figure 3b. The solid lines represent a double exponential fit to the data from Figure 3b with time constants of 5.7 ± 2.4 and 502 ± 28 ms (amplitudes 9 ± 1% and 83 ± 2%, respectively) for the Ile56Thr variant, and with time constants of 6.2 ± 7.1 and 56 ± 7 ms (amplitudes 8 ± 7% and 85 ± 7%, respectively) for wild-type human lysozyme.

along a fast track for all proteins. This represents the fraction of molecules with nativelike protection formed very early in the folding process; such a population of fast-folding molecules has been detected previously in the study of the refolding of both hen and human lysozymes (*3, 8*).

This fast track of folding is particularly clearly visible in the case of the Ile56Thr variant, because the subsequent step is much slower and thus allows a clear distinction between the two kinetic phases. The fast track has a time constant of ca. 6 ms and an amplitude of approximately 10%. For the Ile56Thr variant the main phase has an amplitude of 83 ± 2% and a time constant of 502 ± 28 ms. The main phase of the wild-type protein has an amplitude of 85 ± 7% and a time constant of 56 ± 7 ms, and that of the Asp67His variant an has amplitude of 78 ± 3% and a time constant of 51 ± 6 ms. For all three proteins, the additional 8 ± 3% corresponds to the proportion of protein not protected on the time scale of the pulse labeling measurement and may be associated with the slow phase (phase III) of the fluorescence measurement.

DISCUSSION

Overview of the Refolding Kinetics. Several events take place within 10 ms of initiation of the refolding of wild-type human lysozyme and of the two variants. These include the development of 28−43% of the native fluorescence, hydrogen exchange protection in the A and B helices of ca. 100% of molecules, and the formation of nativelike hydrogen exchange protection in ca. 10% of the molecules. This is attributable to a fast track for folding in human lysozyme as seen previously in this protein (*8*) and in hen lysozyme (*3, 7*). This fast track appears to arise from a population of molecules able to form the native state more efficiently than the remainder of molecules. For the latter, a series of slower steps are involved in the development of the fully native structure (*2*).

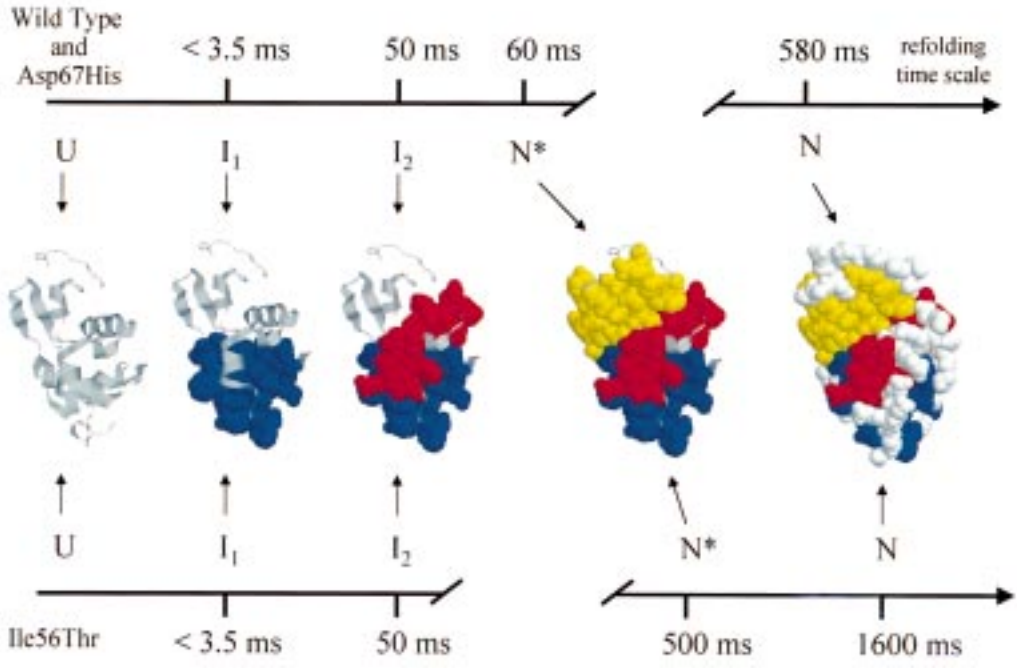
The nature of these steps can be characterized by analysis of the various spectroscopic changes that occur during folding in conjunction with the hydrogen exchange labeling experiments. The fluorescence changes observed early during the stopped-flow measurements (phase I) are likely to be associated with the formation of the first intermediate (I_1) as there are two tryptophans (28 and 34) in helix B. In addition, Trp 112 from helix D has been reported to develop nativelike protection in the dead time of the quenched-flow experiment (*8*). Moreover, the population of molecules folding along the fast pathway (ca. 10%) must give rise to significant changes in the fluorescence signal. It is therefore not possible to attribute unambiguously phase I in the stopped-flow fluorescence experiments to a single event. An important conclusion, however, is that all the fast events contributing to this phase are, within experimental error, the same for both variants and the wild-type protein.

The intermediate kinetic phase observed in the stopped-flow refolding experiments (phase II) is associated with the appearance of nativelike protection in the pulse labeling experiments. The agreement between the time constants determined by stopped-flow and pulse labeling is very good for the wild-type protein (50 ± 10 and 56 ± 7 ms, respectively) and the Asp67His variant (50 and 51 ± 6 ms, respectively). Although the agreement is somewhat less good for the Ile56Thr variant (399 ± 37 and 502 ± 28 ms, respectively), the kinetics are clearly similar. We have seen that the formation of nativelike protection in the majority of molecules occurs via two intermediates, differing in their numbers of hydrogens protected against exchange (Figure 4). Whereas the first intermediate (I_1), having protection in helices A, B, and 3_{10} only, is formed in the dead time of the quenched-flow experiment, it is interesting to note that the formation of the second intermediate having persistent structure in the complete α-domain (I_2) has been reported from NMR data to occur with a time constant of ca. 50 ms in the wild-type protein (*8*). This is in good agreement with the qualitative analysis we have been able to draw from our experiments (Figure 4).

The fact that the formation of I_2 occurs on a time scale similar to that of the native state for the wild-type protein and the Asp67His variant makes it difficult to monitor the intermediate peak in the pulse labeling experiment, as it has only low intensity. By contrast, in the case of the Ile56Thr variant, since the formation of the native state is substantially slower, one can observe much more easily the evolution of the intermediate peak as I_1 converts to I_2. In the pulse labeling experiment with Ile56Thr, the persistence of the intermediate peak can be seen clearly (Figure 3b) and enables a better estimate of the number of protected hydrogens at different folding times (Figure 4). In this regard, the Ile56Thr mutation is a very interesting one, as it results in an uncoupling of the two separate events occurring nearly simultaneously in the wild-type protein, namely, the formation of persistent structure throughout the α-domain and the formation of protection within the β-domain.

Finally, in accord with previous findings (*15, 18, 30*), the kinetics measured by stopped-flow fluorescence show a third, slow, phase in the refolding of wild-type human lysozyme and both variants studied here, of amplitude 5−12%. This phase appears not to be related to proline isomerization in the wild-type protein, as mutation of the proline residues does not suppress this slow phase (*15*). Taniyama et al. have, however, shown that the double mutation Cys77Ala/Cys95Ala produces a variant protein devoid of this slow

Scheme 1: Representation of the Different Steps and Average Time Constants for the Major Refolding Processes Occurring in Wild-Type and Variant Human Lysozymes[a]



[a] The data for Asp67His variant and the wild-type protein are presented together because of the close similarity of their refolding kinetics. N* represents the species which has nativelike protection but has not undergone the final rearrangement step seen in the slow fluorescence phase (III). The nature of this phase is not, however, clear, and it may be that only a small proportion of molecules requires such a rearrangement process. For each species identified along the folding pathway, a structural model generated using the program RasMol (*33*) is shown with the protected regions displayed in space-filled mode colored according to the time of their appearance: blue, <3.5 ms; red, 3.5−50 ms; yellow, >50 ms; white, the remainder.

refolding phase (*30*) and have suggested that the slow refolding phase in the wild-type protein might be due to the existence of two unfolded populations corresponding to two different rotational isomers of the Cys77−Cys95 disulfide bond. Although the kinetics of the slow phase cannot be determined from the pulse labeling experiments, there is ca. 10% of protein not labeled after 2 or even 4 s of refolding; this might be related to the slow phase observed in the fluorescence experiments.

In summary, the data indicate that the folding process results in substantially nativelike structure in the majority of molecules by ca. 60 ms for the wild-type protein and the Asp67His variant and by ca. 500 ms for the Ile56Thr variant. This is illustrated pictorially in Scheme 1, which also brings together the results of the present and previous (*8*) pulse labeling hydrogen exchange data to show the development with time of persistent structure in the wild-type and the variant proteins.

Implications for the Mechanism of Folding. The crystallographic structures of the native states of the Ile56Thr and Asp67His variants of human lysozyme are closely similar to that of the wild-type protein (*17*). This is particularly evident for the Ile56Thr variant; despite the introduction of a hydrophilic residue into the interior of the protein, the root-mean-square deviation between the two structures for all residues is below 1 Å. The Asp67His variant shows greater differences in local regions from the wild-type structure. The conformations of two external loops in the β-domain are significantly perturbed (rmsd as high as 9.85 Å). This is attributable to the loss of hydrogen bonds between Asp67 and Tyr54, between Asp67 and Lys69, and between Asp67 and Thr70 (Figure 5). The observation that the fastest steps in refolding (phase I) are closely similar in the different proteins indicates that the events associated with the fast changes in fluorescence are not substantially altered by the destabilizing effects of either mutation. One of the events occurring in this phase is the formation of persistent structure in helices A and B of the α-domain, as shown by the hydrogen exchange pulse labeling experiments (*8*). That this specific process is not detectably perturbed by either mutation is consistent with the fact that neither of the mutated residues is in or near this region.

Phase II of the folding process has been associated with the subsequent formation of the native structure by coalescence of first the helices C and D, and then the β-domain, onto the helices A and B (Scheme 1). The lack of effect of the Asp67His mutation on these steps supports this conclusion because this residue is not in the interface region of either of these segments of structure. The lack of perturbation of the kinetics also suggests that this residue is not involved in nativelike structure in the transition state ensemble of this folding step (*31*). In accord with this conclusion, this mutation has a large effect on the unfolding rate of the protein as the structural changes destabilize the native state relative to the transition state of the unfolding reaction.

In contrast to the behavior of the Asp67His variant, phase II of folding is substantially slower in the Ile56Thr variant. This suggests that the residue is in a structured region of the transition state for folding. Its location at the interface of the α- and β-domains is again fully supportive that this step of folding involves the generation of nativelike structure at this domain interface region. In addition, the much smaller

effect of this mutation on the kinetics of the unfolding step compared to the effect of the Asp67His mutation indicates that this residue is in a region partially structured in the transition state for unfolding.

In conclusion, the results presented here highlight the critical nature of the location of the mutation in its effect on either the unfolding kinetics (as in the case of the Asp67His variant) or both the unfolding and refolding kinetics (as in the case of the Ile56Thr variant). Identification of which refolding steps are specifically affected by the mutations has enabled information about their structural nature to be obtained. The results are fully consistent with the structural conclusions drawn from the pulse labeling and optical spectroscopic measurements.

Implications for the Mechanism of Amyloid Formation. The native states of the two amyloidogenic variants of lysozyme and that of the wild-type protein have closely similar structures. Indeed, both variants are enzymatically active with K_M and k_{cat} reasonably close to those of the wild-type protein (*17*). In addition, the present data show that the unfolding and refolding processes of both variants in vitro are fundamentally similar to those of the wild-type protein. There are, however, variations in the rates of folding and unfolding, and it is clear that the effects of the mutations are very different for the two variants. Interestingly, however, despite these differences in kinetics, the overall destabilization of the structure relative to the wild-type protein is similar in the two variants. Thus, the temperatures of the midpoints of thermal unfolding of the variant proteins differ from the wild-type protein by 12 °C in both cases (*17*). In addition, although the Asp67His variant unfolds ca. 6 times faster than the Ile56Thr variant in 5.4 M GuHCl, it refolds faster in 0.54 M GuHCl by almost the same factor. This suggests a similar overall equilibrium constant for the folding reaction for the two variants. The relatively low level of destabilization should permit the variant proteins to function in a manner similar to that of the wild-type protein under normal physiological conditions. However, even under conditions in which the native states are stable, the transient population of partially folded intermediates will be increased very substantially for both variants relative to the wild-type protein. As such species will expose to solvent at least some residues normally buried in the close packing interior of the protein, they are likely to be prone to aggregation.

In the case of both lysozyme variants, the β-domain is destabilized relative to the α-domain, suggesting that this region could be particularly important for the development of intermolecular interactions. As this region of the protein forms a β-sheet in the native protein, initial steps in the aggregation process could involve the formation of intermolecular rather than intramolecular hydrogen bonds. Later steps, however, will undoubtedly require substantial rearrangements in the aggregated species to form the characteristic fibrils observed in electron micrographs. The partially folded intermediate structure may, however, be important for rapid and effective formation of the initial intermolecular contacts, even though its structure is unlikely to be a significant factor in defining the overall conformation of the fibrils themselves. These appear to be closely similar for a wide range of different proteins regardless of the structural propensities of the polypeptides that gives rise to them (*32*).

Given the very low concentrations of intermediates likely under normal physiological conditions [even for the amyloidogenic variants only ca. 1 molecule in 1000 of lysozyme will be in this state at pH 5.0 and 37 °C, according to hydrogen exchange data in near EX2 conditions (*17*)], amyloid formation in vivo is expected to be very slow. It is likely to be enhanced, however, by localization of the proteins in compartments under other conditions where stability is significantly reduced, e.g., by exposure to low pH or a membrane surface and also by seeding, either in situ or from an external source (*20*). The existence of this complex balance of conditions may be the reason that for only a few proteins does amyloid formation appear to be able to compete with either refolding to the native state or degradation in a physiological environment. The lysozyme variants represent examples of mutations which allow slow amyloid formation to occur despite the intrinsic control mechanisms which mitigate against such a process in vivo. The fact that these variants have significantly different folding and unfolding kinetics, but very similar equilibrium stability, is particularly interesting. It suggests that the overall stability, rather than the specific nature of the folding or unfolding kinetics, is a crucial factor in amyloidogenicity. The lysozyme variants are likely to represent proteins which are sufficiently stable to be expressed and functional under normal circumstances but unstable enough to allow sufficiently high concentrations of aggregation-prone intermediates to accumulate under aberrant conditions in tissues where slow aggregation can take place.

ACKNOWLEDGMENT

We thank Professor M. B. Pepys (Hammersmith Hospital, London) for valuable discussions and continuing interest in this work.

REFERENCES

1. Dobson, C. M., Evans, P. A., and Radford, S. E. (1994) *Trends Biochem. Sci. 19*, 31.
2. Radford, S. E., and Dobson, C. M. (1995) *Philos. Trans. R. Soc. London B 348*, 17.
3. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature 358*, 302.
4. Chaffotte, A. F., Guillou, Y., and Goldberg, M. E. (1992) *Biochemistry 31*, 9694.
5. Itzhaki, L. S., Evans, P. A., Dobson, C. M., and Radford, S. E. (1994) *Biochemistry 33*, 5212.
6. Matagne, A., Radford, S. E., and Dobson, C. M. (1997) *J. Mol. Biol. 267*, 1068.
7. Wildegger, G., and Kiefhaber, T. (1997) *J. Mol. Biol. 270*, 294.
8. Hooke, S. D., Radford, S. E., and Dobson, C. M. (1994) *Biochemistry 33*, 5867.
9. Hore, P. J., Winder, S. L., Roberts, C. H., and Dobson, C. M. (1997) *J. Am. Chem. Soc. 119*, 5049.
10. Miranker, A. D., Robinson, C. V., Radford, S. E., Aplin, R. T., and Dobson, C. M. (1993) *Science 262*, 896.
11. Hooke, S. D., Eyles, S. J., Miranker, A., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1995) *J. Am. Chem. Soc. 117*, 7548.
12. Matagne, A., Chung, E. W., Ball, L. J., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1998) *J. Mol. Biol. 277*, 997.
13. Kulkarni, S. K., Ashcroft, A. E., Carey, M., Masselos, D., Robinson, C. V., and Radford, S. E. (1999) *Protein Sci. 8*, 35.

14. Kato, S., Shimamoto, N., and Utiyama, H. (1982) *Biochemistry 21*, 38.
15. Herning, T., Yutani, K., Taniyama, Y., and Kikuchi, M. (1991) *Biochemistry 30*, 9882.
16. Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M., and Hsuan, J. J. (1993) *Nature 362*, 553.
17. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997) *Nature 385*, 787.
18. Funahashi, J., Takano, K., Ogasahara, K., Yamagata, Y., and Yutani, K. (1996) *J. Biochem. 120*, 1216.
19. Taniyama, Y., Yamamoto, Y., Nakao, M., Kikuchi, M., and Ikehara, M. (1988) *Biochem. Biophys. Res. Commun. 152*, 962.
20. Kelly, J. W. (1998) *Curr. Opin. Struct. Biol. 8*, 101.
21. Prodromou, C., and Pearl, L. H. (1992) *Protein Eng. 5*, 827.
22. Archer, D. B., Jeenes, D. J., and Mackenzie, D. A. (1994) *Antonie Van Leeuwenhoek 65*, 245.
23. van Hartingsveldt, W., Mattern, I. E., van Zeijl, C. M., Pouwels, P. H., and van den Hondel, C. A. (1987) *Mol. Gen. Genet. 206*, 71.
24. Archer, D. B., MacKenzie, D. A., Ridout, M. J., Kikuchi, M., Yamamoto, Y., Taniyama, Y., Ishimaru, K., Yoshikawa, W., Kaisho, Y., and Ikehara, M. (1995) *Appl. Microbiol. Biotechnol. 44*, 157.
25. Kikuchi, M., Yamamoto, Y., Taniyama, Y., Ishimaru, K., Yoshikawa, W., Kaisho, Y., and Ikehara, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9411.
26. MacKenzie, D. A., Spencer, J. A., Le Gal Coeffet, M. F., and Archer, D. B. (1996) *J. Biotechnol. 46*, 85.
27. Pace, C. N. (1986) *Methods Enzymol. 131*, 266.
28. Wilm, M., and Mann, M. (1996) *Anal. Chem. 68*, 1.
29. Redfield, C., and Dobson, C. M. (1990) *Biochemistry 29*, 7201.
30. Taniyama, Y., Ogasahara, K., Yutani, K., and Kikuchi, M. (1992) *J. Biol. Chem. 267*, 4619.
31. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) *J. Mol. Biol. 254*, 260.
32. Sunde, M., and Blake, C. C. F. (1998) *Q. Rev. Biophys. 31*, 1.
33. Sayle, R. A., and Milner-White, E. J. (1995) *Trends Biochem. Sci. 20*, 374.
34. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946.
35. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505.
36. Bacon, D. J., and Anderson, W. F. (1988) *J. Mol. Graphics 6*, 219.
37. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869.

BI983037T