# Kinetic Consequences of the Removal of a Disulfide Bridge on the Folding of Hen Lysozyme†

Stephen J. Eyles,‡ Sheena E. Radford,‡ Carol V. Robinson,§ and Christopher M. Dobson*,‡

*New Chemistry Laboratory and Dyson Perrins Laboratory, Oxford Centre for Molecular Sciences, South Parks Road, Oxford, OX1 3QT, U.K.*

*Received June 27, 1994; Revised Manuscript Received August 19, 1994®*

ABSTRACT: Quenched-flow hydrogen exchange labeling, monitored by $^1$H NMR and electrospray ionization mass spectrometry (ESI-MS), has been employed in conjunction with stopped-flow circular dichroism and fluorescence to study the kinetic refolding from guanidinium chloride of a derivative of hen lysozyme in which one of the four disulfide linkages (Cys6–Cys127) has been selectively chemically reduced and carboxymethylated ($CM^{6,127}$-lysozyme). Removal of this disulfide bridge has little effect on the structure and activity of the native enzyme, and the overall kinetics of refolding are very similar to those of the unmodified protein. A substantial amount of secondary structure is formed within 2 ms of the initiation of folding, followed by the slower formation of tertiary interactions characteristic of the native state, which are attained with a time constant ($\tau$) of *ca.* 200 ms. There is clear evidence for fast and slow refolding populations, as in the intact protein. Folding of the three-disulfide derivative does, however, exhibit a major difference from that of the intact protein under the same final refolding conditions, in that the transient intermediate on the major refolding pathway of the intact protein, having persistent structure in the $\alpha$-helical domain of the protein, is not detected by hydrogen exchange labeling during folding of the three-disulfide derivative. This suggests that the disulfide bond linking the N- and C-terminal regions of the protein is crucial for stabilization of the partially folded intermediate. In addition, the overshoot in the far-UV CD and the fluorescence minimum, both of which are attributed to non-native interactions, is not observed in the folding of $CM^{6,127}$-lysozyme. That the lack of a detectable stable intermediate in the folding of $CM^{6,127}$-lysozyme does not significantly affect the rate of attainment of the native state of the protein supports the proposed independent nature of the two folding domains and, as the Cys6–Cys127 disulfide bond is located in the $\alpha$-domain, indicates that the rate-limiting step in folding of the intact protein, as well as of the three-disulfide derivative, involves stabilization of the $\beta$-domain. The role of disulfide bridges in the formation and maintenance of the three-dimensional fold of proteins and in facilitating the observation of marginally stable intermediate species is discussed.

Intensive efforts have been made over a number of years to elucidate the folding pathways of proteins. Recent advances in instrumentation have allowed a number of complementary techniques to be applied to mechanistic and kinetic studies of the folding of small globular proteins (Evans & Radford, 1994). One method is to trap intermediates as they are formed, for instance, by blocking cysteine residues during the oxidative refolding of disulfide-containing proteins (Creighton, 1992; Weissman & Kim, 1992). However, due to the transient nature of many intermediate species involved in the formation of noncovalent secondary and tertiary interactions, alternative methods of detection are often necessary. Elegant experiments using pulsed hydrogen exchange labeling methods in conjunction with NMR[1] have enabled the detailed structural characterization of such transiently formed species (Roder *et al.*, 1988; Udgaonkar & Baldwin, 1988; Radford *et al.*, 1992b). More recently, the use of electrospray ionization mass spectrometry (ESI-MS) has been used to give information on the populations of species present during the refolding of hen lysozyme (Miranker *et al.*, 1993). In conjunction with these techniques, rapid mixing optical methods have enabled information to be gained about global folding characteristics (Ikeguchi *et al.*, 1986; Ptitsyn *et al.*, 1990; Chaffotte *et al.*, 1992; Elöve *et al.*, 1992; Evans & Radford, 1994; Itzhaki *et al.*, 1994). Structural studies of stable, partially folded species have also been important in giving insight into the nature of folding intermediates (Baum *et al.*, 1989; Harding *et al.*, 1991; Buck *et al.*, 1993; Fan *et al.*, 1993), and site-directed mutagenesis has been employed to study the effects of specific interactions during folding (Fersht, 1993; Fersht & Serrano, 1993; Matthews, 1993).

Several common features have emerged from these studies, particularly that compact collapsed states with a significant amount of secondary structure content form very rapidly for a number of proteins (Kuwajima *et al.*, 1985; Elöve *et al.*, 1992; Jennings & Wright, 1993; Khorasanizadeh *et al.*, 1993; Kuwajima *et al.*, 1993; Varley *et al.*, 1993). These species have at least some characteristics of partially folded states observed at equilibrium for a number of proteins under mildly denaturing conditions (Baum *et al.*, 1989; Hughson *et al.*, 1990; Harding *et al.*, 1991; Buck *et al.*, 1993). Similarly,

† This work is a contribution from the Oxford Centre for Molecular Sciences, funded jointly by the BBSRC, the EPSRC, and the MRC. S.E.R. is a Royal Society (1983) University Research Fellow. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Institute.

* Author to whom correspondence should be addressed.

‡ New Chemistry Laboratory.

§ Dyson Perrins Laboratory.

® Abstract published in *Advance ACS Abstracts,* October 1, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; ESI-MS, electrospray ionization mass spectrometry; CD, circular dichroism; HEWL, hen egg white lysozyme; GuHCl, guanidinium chloride; COSY, *J*-correlated spectroscopy; $CM^{6,127}$-lysozyme, carboxymethyl-(Cys6–Cys127)-lysozyme; $T_m$, midpoint temperature of thermal denaturation; $C_m$, midpoint denaturant concentration; Da, dalton.

0006-2960/94/0433-13038$04.50/0 

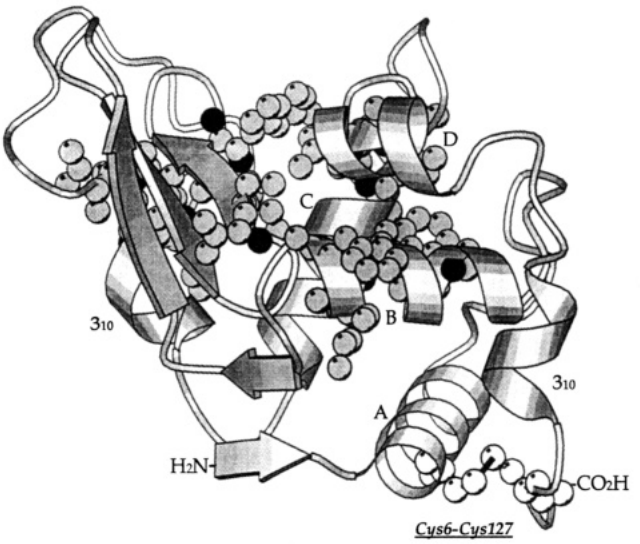


FIGURE 1: Schematic diagram of the native structure of hen lysozyme (Blake, 1965). The Cys6–Cys127 disulfide bridge is shown in white. The amino acid residues whose backbone amides are most highly protected against hydrogen exchange in the native state of $CM^{6,127}$-lysozyme at pH 3.8, 20 °C, are indicated in gray, with the corresponding amide nitrogen represented in black. This figure was drawn using the program Molscript (Kraulis, 1991).

pulse labeling experiments coupled with stopped-flow circular dichroism in the far UV have shown that helical regions and simple motifs are often formed faster than more complex structures such as β-sheets (Radford *et al.*, 1992b; Varley *et al.*, 1993).

The kinetic folding of hen egg white lysozyme has been studied extensively using a variety of different techniques (Radford *et al.*, 1992b; Dobson *et al.*, 1994). The first detectable step (≤2 ms) involves very rapid formation of a collapsed state containing extensive secondary structure, although this structure clearly is not sufficiently stable to afford substantial protection from amide exchange. A majority of molecules in the refolding mixture passes through an intermediate species in which the four α-helices and a C-terminal $3_{10}$-helix (the α-domain, comprising residues 1–36 and 87–129) form persistent structure cooperatively, substantially protecting (protection factor ≥ 500) the amides in this domain from exchange. In a slower step, the remainder of the protein (the β-domain, residues 37–86), including two β-sheet regions, a long loop and a $3_{10}$-helix, folds to form the native structure. The overall folding process, under the conditions employed for these experiments, is more complex, however, since parallel pathways are involved; a proportion of the molecules becomes substantially protected in both the α- and β-domains, forming apparently native-like structure, on a faster time scale (Radford *et al.*, 1992b; Miranker *et al.*, 1993).

Lysozyme has four disulfide bridges, which remain intact in the studies of the unmodified protein described in this paper (Figure 1). Two of these (Cys6–Cys127 and Cys30–Cys115) are located in the α-domain: one (Cys64–Cys80) within the β-domain and one (Cys76–Cys94) linking the two domains. Recent studies on the effects of disulfide bridges on folding rates have indicated that the introduction of disulfides can dramatically alter the kinetics of folding. For example, an 850-fold increase in the refolding rate was observed for a modified form of subtilisin BPN′, in which a disulfide was introduced (Strausberg *et al.*, 1993), and was explained in terms of a decrease in configurational entropy in the transition state for folding. Disulfide bridges introduced into barnase have also been shown to affect folding kinetics by stabilizing either the intermediate or transition state of folding, depending on their precise location in the structure (Clarke & Fersht, 1993). In this paper, we describe detailed kinetic studies of the folding of a derivative of hen lysozyme, in which one of the disulfide bridges located in the α-domain, Cys6–Cys127, has been selectively removed by chemical modification; the three remaining disulfide bonds remain intact in their native pairings.

We have previously shown, by differential scanning calorimetry, that the decreased thermal stability of this three-disulfide derivative with respect to the intact protein is primarily an entropic effect, as a result of increased disorder in the denatured state in the absence of the constraining disulfide bridge (Cooper *et al.*, 1992). This linkage is of particular interest for the folding of lysozyme, since it links the N- and C-terminal regions in the α-domain of the protein. Using hydrogen exchange pulse labeling techniques, in conjunction with stopped-flow circular dichroism and intrinsic tryptophan fluorescence, we have investigated in detail the folding of this three-disulfide derivative. By comparison with the extensive data available for the unmodified protein, we discuss the importance of this disulfide bridge in the folding pathway of hen lysozyme and, more generally, the importance of disulfide bonds in protein folding.

## MATERIALS AND METHODS

$CM^{6,127}$-lysozyme was prepared and purified extensively by a modification of the method previously described (Radford *et al.*, 1991). Hen lysozyme (20 mg/mL) was incubated with constant mixing at 15 °C with 5 mM dithiothreitol in 100 mM Tris–acetate buffer at pH 7.8. The instant turbidity induced by the formation of insoluble aggregates was monitored by light scattering at 600 nm. When this turbidity had increased 40-fold relative to the starting point, the reaction was quenched by the addition of 25 mM iodoacetic acid and incubated for 15 min at pH 8.0. Further reactions, such as the carboxymethylation of His15 (Yamada *et al.*, 1984) or oxidations caused by the presence of free iodine (Imoto & Rupley, 1973), were avoided by adjusting the pH to 3 with HCl, followed by extensive dialysis of the protein at 4 °C against 1 mM HCl. The protein was then lyophilized. The lyophilized protein was suspended in a minimum volume of $H_2O$, and the insoluble fraction was removed by centrifugation (3000*g* for 15 min at 4 °C). $CM^{6,127}$-lysozyme was then purified by ion exchange chromatography on a Bio-Rex 70 column (Bio-Rad) and eluted with 0.15 M sodium phosphate buffer (pH 7.0). Unmodified lysozyme that had bound to the column was then eluted using the same buffer containing 1 M NaCl. Purified $CM^{6,127}$-lysozyme was dialyzed extensively at 4 °C against 0.16 mM HCl (pH 3.8) and then lyophilized. Extreme care was taken to ensure the highest purity of the $CM^{6,127}$-lysozyme used in refolding studies, since any contaminant of unmodified lysozyme could mask refolding of the derivative in the optical measurements. Purity was confirmed by acid urea polyacrylamide gel electrophoresis, amino acid analysis (Moore *et al.*, 1968; Hendrikson & Meredith, 1984), and ESI-MS and was found to be >98% pure. The total average yield of pure $CM^{6,127}$-lysozyme was ≤10% of the starting material.

*Pulsed Amide Hydrogen Exchange Labeling.* Quenched-flow experiments were carried out with a Bio-Logic QFM-5 module, using a modification of the protocol described previously (Radford *et al.*, 1992b). Deuterated $CM^{6,127}$-lysozyme, in which all of the labile NH hydrogens were

exchanged for deuterium, was prepared by incubation in $D_2O$ at pH 3.8, 60 °C, for 10 min, followed by lyophilization. This process was repeated three times to ensure complete deuteration of all exchangeable amides. $CM^{6,127}$-lysozyme was denatured at 20 mg/mL in 3 M GuDCl. Refolding was initiated by 11-fold dilution into 20 mM sodium acetate buffered to pH 4.5. After various refolding times between 3.5 ms and 5 s, all unprotected amide deuterons were exchanged for protons by the addition of a volume of 20 mM sodium borate buffer (pH 9.5) that was 5 times that of the starting volume. After a labeling time of 8.4 ms, exchange was quenched by reducing the pH to 4.0 by further dilution with 0.5 M acetic acid. The protein was then allowed to refold completely under these conditions where further hydrogen exchange was negligible. Samples were concentrated using Centriprep-10 and Centricon-10 microconcentrators (Amicon, Inc.). For NMR experiments, samples were buffer-exchanged into 40 mM deuterated sodium acetate at pH 3.8 to minimize the need for solvent suppression. The final concentration of samples for NMR experiments was approximately 2 mM.

As a control, a sample of protein was heated to 80 °C for 15 min in the same final buffer and isotope concentrations as for the labeling pulse to ensure the complete labeling of all amides with correct isotopic distribution. The sample was buffer-exchanged in the same way as previously described. This sample was assumed to have 100% proton occupancy and was used to normalize the peak intensities for each amide. A further control was carried out by using an initial protein concentration of 2 mg/mL to ensure no dependence of refolding on the protein concentration. Several time points were measured at each concentration and were confirmed to be identical, within error, at the two concentrations examined. For pulse intensity experiments, the pH of the borate buffer was adjusted to give the desired labeling pH between pH 9.0 and 10.5, and the concentration of acetic acid was varied accordingly to ensure efficient quenching to pH 4.

*NMR Spectroscopy.* Phase-sensitive *J*-correlated spectroscopy (Aue *et al.*, 1976; Bax & Freeman, 1981) was performed at 500 MHz, on a GE/Nicolet spectrometer belonging to the Oxford Centre for Molecular Sciences, at a temperature of 20 °C. This temperature was chosen to minimize amide hydrogen exchange during acquisition while maximizing resolution of the cross peaks. Data sets consisting of 256 complex $t_1$ increments of 32 transients were collected, using a sweep width of 7042 Hz and a relaxation delay of 700 ms. Data were processed and peak-picked using the FTNMR and FELIX programs (Hare Research Inc.) on a SUN workstation. Resolution enhancement was by trapezoidal multiplication and double-exponential multiplication in $t_2$ and by trapezoidal multiplication alone in $t_1$. After zero-filling, the final digital resolution was 1.7 Hz/point in both dimensions.

$\alpha$CH–NH cross-peak intensities were measured as the sum of absolute peak heights of the four phase-sensitive components of each peak, and intensities in each spectrum were normalized to the aromatic cross peaks of Tyr23 and Tyr53. Individual cross-peak intensities were then normalized to the values obtained from the zero time point sample. Protection profiles were fitted to the sum of two exponential functions by nonlinear regression analysis, using the program Kaleidagraph (Abelbeck Software, Inc.).

*Electrospray Ionization Mass Spectrometry.* For mass spectrometry experiments, samples from pulse labeling experiments were exhaustively desalted by washing on Centricon-10 membranes with Milli-Q water adjusted to pH 3.8 with HPLC-grade formic acid. Experiments were performed using a modification of the method of Miranker *et al.* (1993), using a lower temperature and a less stringent solvent system to minimize unfolding and hydrogen exchange of this less stable protein. All mass spectra were acquired on a platform single analyzer mass spectrometer equipped with an electrospray interface (VG Biotech, Fisons Intruments). The sample interface was equilibrated overnight with water/formic acid at 4 °C and pH 3.8. The solvent delivery system (Michrom BioResources, Inc.) and the nebulizer gas were immersed in an ice/salt bath. Protein samples at 20 pmol/$\mu$L in water/formic acid (pH 3.8) were introduced into the ion source at 10 $\mu$L/min using the same buffer as the mobile phase. The instrument was scanned from 1260 to 1480 $m/z$ and calibrated against protonated unmodified hen lysozyme. Under these conditions, no hydrogen exchange was detected during the acquisition time of the spectrum ($\sim$2 min).

*Circular Dichroism.* To monitor equilibrium unfolding of $CM^{6,127}$-lysozyme, samples were dissolved to a protein concentration of 0.2 mg $mL^{-1}$ in unbuffered GuHCl solutions adjusted to pH 3.8 with small amounts of HCl or NaOH. Ellipticity at 222 nm was measured on a Jasco J720 spectropolarimeter in 1 mm cells. Stopped-flow experiments were carried out using a Jasco J720 spectropolarimeter equipped with a Bio-Logic SFM-3 stopped-flow module. $CM^{6,127}$-lysozyme at 20 mg/mL for near-UV CD or 2 mg/mL for far-UV CD was dissolved in 3 M GuHCl. Refolding was initiated by an 11-fold dilution into 20 mM sodium acetate buffered to pH 4.5, thus giving refolding conditions identical to those of the pulse labeling experiments. Refolding was monitored over a period of 3 s at either 289 nm (near UV) or 225 nm (far UV) using a 2 ms step resolution, averaging data points over 4 ms.

*Fluorescence.* Experiments were carried out using refolding conditions identical to those of the near-UV CD measurements. An Applied Photophysics SX17MV stopped-flow instrument, equipped with 2.5 mL and 250 $\mu$L syringes to allow a 10:1 (buffer/protein) mix, was used with an excitation wavelength of 285 nm. Total fluorescence emission above 320 nm was measured using a high-pass filter. Refolding was monitored for 2 s with a 1 ms step resolution.

Formation of the active site was monitored by binding of the fluorescently labeled inhibitor (4-methylumbelliferyl)-*N*,*N*′-diacetylglucosamine (MeU-diNAG), which was added to a concentration of 6 $\mu$M in the refolding buffer. Total fluorescence above 335 nm was monitored, with excitation at 330 nm.

## RESULTS

*Hydrogen Exchange in Native $CM^{6,127}$-Lysozyme.* The $^1$H NMR spectra of lysozyme and $CM^{6,127}$-lysozyme have been completely assigned at 35 °C and pH 3.8 (Redfield & Dobson, 1988; Radford *et al.*, 1991). Both proteins are native under these conditions with virtually identical structures, as judged by NMR and X-ray crystallography (Radford *et al.*, 1991; Hill *et al.*, 1993) and by optical methods (Denton & Scheraga, 1991); indeed, the three-disulfide derivative largely retains its enzymatic activity (Radford *et al.*, 1991). At this temperature, however, although $CM^{6,127}$-lysozyme is still folded, the majority of backbone amides exchange, some to a substantial degree, with bulk solvent during the 4 h acquisition time of a COSY experiment, whereas amides in the unmodified protein are stable against exchange for weeks under these conditions (Pedersen *et al.*, 1991). One possible explanation for this is that $CM^{6,127}$-lysozyme has a lower $T_m$

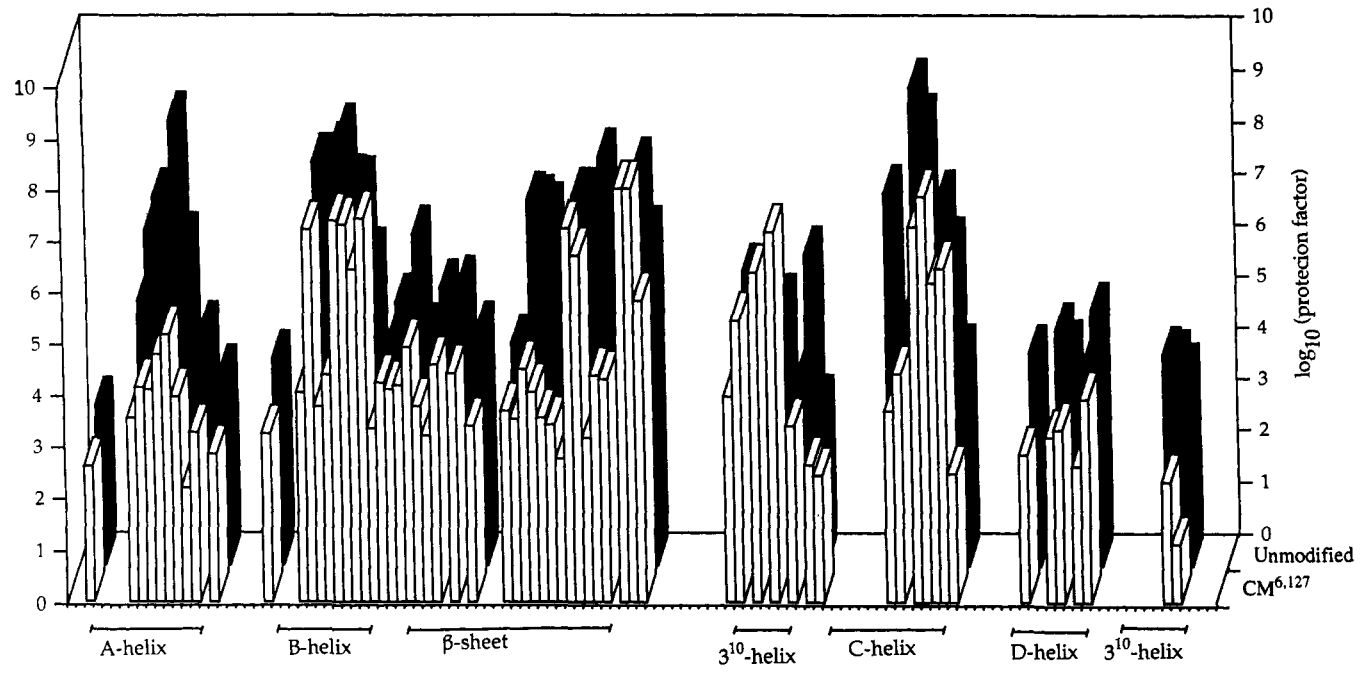


FIGURE 2: Amide hydrogen exchange protection in the native state of $CM^{6,127}$-lysozyme, pH 3.8, 20 °C (shown in white), and unmodified hen lysozyme, pH 7.5, 30 °C (black). $\log_{10}$(protection factor) is plotted *versus* the sequence for the two proteins.

(53 °C at pH 3.8) than that of the unmodified protein (77 °C at pH 3.8) (Wedin *et al.*, 1982), thus allowing increased exchange *via* an unfolding mechanism (Kim & Woodward, 1993). In accord with this, at a lower temperature of 20 °C, hydrogen exchange is significantly retarded in native $CM^{6,127}$-lysozyme, allowing exchange to be monitored readily by 2D NMR. Under these conditions, exchange in the native protein is insignificant during spectral acquisition, while the increased line widths and cross-peak overlap at the lower temperature do not reduce substantially the number of amides that can be followed. Spectral assignment at 20 °C was readily possible by direct comparison with spectra at 35 °C, since all changes in αCH chemical shifts were relatively small (≤0.1 ppm). Assignments were confirmed by means of NOESY experiments.

The rates of exchange in the native state of $CM^{6,127}$-lysozyme at 20 °C were monitored for 68 amides. Protection factors, defined as the ratio of the intrinsic rate of hydrogen exchange in an unstructured polypeptide to the measured hydrogen exchange rate in the protein under identical conditions, were calculated using the latest correction factors for neighboring side chain effects (Bai *et al.*, 1993). Figure 2 shows the pattern of protection for $CM^{6,127}$-lysozyme (pH 3.8, 20 °C) and beside it the corresponding data obtained previously for the unmodified protein (pH 7.5, 30 °C) (Radford *et al.*, 1992a), recalculated using the new correction factors of Bai *et al.* (1993). The overall pattern of protection in $CM^{6,127}$-lysozyme is similar to that of the intact protein, in that those amides that are most highly protected from exchange lie in regions of well-defined secondary structure, confirming the structural integrity of $CM^{6,127}$-lysozyme. The calculated protection factors are, however, generally lower in the case of the three-disulfide derivative. The most noticeable difference is that protection from exchange is, on average, approximately 100-fold lower in the N- and C-terminal regions of $CM^{6,127}$-lysozyme, near the site of modification. The central strand of the triple-stranded β-sheet is also somewhat less protected in $CM^{6,127}$-lysozyme, while in other regions the differences are relatively small.

The amides monitored in $CM^{6,127}$-lysozyme can be divided into two distinct classes, namely, those with protection factors $\leq 10^5$ and those with protection factors exceeding $10^5$. The more highly protected class consists of 17 residues, the majority of which are clustered in a region far from the site of modification: nine residues in the B- and C-helices of the protein, two amides in the third strand of the triple-stranded β-sheet, and six in the long loop region, as indicated in Figure 1. These residues are quite close in the tertiary structure in a region containing two disulfide bridge cross-links (64–80 and 76–94), and there is good correspondence with the most highly protected amides in unmodified lysozyme, which have protection factors of $\sim 10^8$ (Pedersen *et al.*, 1991). By contrast, in the unmodified protein a large number of other amides, found throughout the entire protein sequence have protection factors exceeding $10^5$. A calculation using the methods of Lee and Richards (1971) indicates very little difference in solvent accessibility between the native states of the unmodified protein and its three-disulfide derivative. This suggests that a stable hydrogen exchange core persists within the structure of $CM^{6,127}$-lysozyme that is of slightly lower stability than that of the unmodified protein, whereas the remainder of the molecule is less stable against hydrogen exchange, presumably due to the increased frequency of local unfolding events following removal of the Cys6–Cys127 disulfide bridge. Nevertheless, the protection factors in the three-disulfide protein are sufficiently high to enable pulsed hydrogen exchange labeling experiments measured by NMR to be performed at 20 °C.

*Hydrogen Exchange Pulse Labeling Experiments Monitored by NMR.* The thermal denaturation of $CM^{6,127}$-lysozyme has been shown previously by optical methods and by NMR to be highly cooperative (Denton & Scheraga, 1991; Radford *et al.*, 1991). The unfolding of $CM^{6,127}$-lysozyme by guanidinium chloride at pH 3.8 was monitored in the present work by far-UV circular dichroism and found, like the unmodified protein, also to be a cooperative process under these conditions. However, consistent with the lower stability of $CM^{6,127}$-lysozyme, the midpoint of the transition was

Table 1: Comparison of Folding Rates and Kinetic Amplitudes of Unmodified Hen Lysozyme and $CM^{6,127}$-Lysozyme[a]

| protein refolding conditions/experiment | unmodified,[b] 6 M→ 0.55 M GuHCl, pH 5.5 | | unmodified, 6 M→ 0.27 M GuHCl, pH 4.5 | | $CM^{6,127}$, 3 M→ 0.27 M GuHCl, pH 4.5 | |
|---|---|---|---|---|---|---|
| CD (225 nm) | 80% | <2 ms | 80% | <2 ms | 80% | <2 ms |
| | 30% | 25 ± 12 ms | 30% | 15 ± 8 ms | | |
| | −10% | 345 ± 70 ms | −10% | 150 ± 30 ms | 20% | 200 ± 20 ms |
| CD (289 nm) | 100% | 330 ± 60 ms | 100% | 235 ± 30 ms | 100% | 205 ± 20 ms |
| Trp fluorescence | 120% | <1 ms | 120% | <1 ms | 80% | <1 ms |
| | 290% | 25 ± 3 ms | 290% | 20 ± 2 ms | | |
| | −310% | 340 ± 20 ms | −310% | 190 ± 20 ms | 20% | 267 ± 15 ms |
| MeU-diNAG fluorescence | 100% | 350 ± 50 ms | 100% | 210 ± 30 ms | 100% | 230 ± 30 ms |
| NMR | | | | | | |
| α-domain | 48% | 5 ± 3 ms | 45% | 6 ± 4 ms | 34% | 13 ± 4 ms |
| | 52% | 65 ± 25 ms | 55% | 50 ± 30 ms | 66% | 278 ± 140 ms |
| β-domain | 29% | 11 ± 5 ms | 30% | 10 ± 5 ms | 35% | 19 ± 6 ms |
| | 71% | 340 ± 180 ms | 70% | 230 ± 110 ms | 65% | 347 ± 180 ms |
| ESI-MS | | nd[c] | | nd | 31% | 16 ± 5 ms |
| | | | | | 69% | 300 ± 30 ms |

[a] Errors indicated represent not only the error of the fit but also the average of several data sets. [b] Results from previously published data are shown for comparison (Radford *et al.*, 1992b; Itzhaki *et al.*, 1994). [c] nd, not determined.

reduced from 4.5 M GuHCl for wild-type lysozyme to 1.5 M for the three-disulfide derivative. From these data, an initial denaturant concentration of 3 M GuHCl was chosen for the following refolding experiments to ensure that $CM^{6,127}$-lysozyme was completely unfolded and yet to retain a sufficiently low final concentration of GuHCl (0.27 M) after 11-fold dilution to allow complete refolding of the protein and to maximize the possibility of detection of any transient, marginally stable intermediates.

In order to make a direct comparison of the folding of $CM^{6,127}$-lysozyme with that of the intact protein, it is vital that the final denaturant conditions employed for refolding are identical (Kato *et al.*, 1981; M. Kotik, S. J. Eyles, S. E. Radford, and C. M. Dobson, unpublished results). Control experiments were performed with each of the techniques that will be described later, in which unmodified lysozyme, denatured in 6 M GuHCl, was refolded by 21-fold dilution into sodium acetate buffer (pH 4.5) in order to achieve a final denaturant concentration of 0.27 M, identical to the conditions employed for the refolding studies on $CM^{6,127}$-lysozyme. Refolding rates and kinetic amplitudes for unmodified lysozyme refolded under these conditions are compared in Table 1. No significant differences were detected as a result of refolding at pH 4.5 rather than pH 5.5 as previously published (Radford *et al.*, 1992b), which is consistent with there being a few ionizable groups having $pK_a$'s within this pH range.

Quenched-flow experiments were performed to probe the kinetic refolding of $CM^{6,127}$-lysozyme at 20 °C and pH 4.5, where the native protein is stable and hydrogen exchange in the refolding molecules is sufficiently slow compared with the time course studied. The kinetics of amide protection during refolding could thus be measured unambiguously for 38 backbone amides. A control in which native, deuterated $CM^{6,127}$-lysozyme was exposed to an 8.4 ms labeling pulse at pH 9.5, followed by the usual quenching and washing steps, confirmed that labeling of the native protein by such a pulse is negligible for the 38 amides monitored.

Kinetic profiles for a representative selection of individual amides are shown in Figure 3. All of the data fit to a double-exponential process: 20–30% of the amplitudes have a time constant of ~20 ms, with a slower phase ($\tau \sim 300$ ms) accounting for an additional 50–60% of the total protection. These amplitudes and rates are similar for every measured amide, suggesting a high degree of cooperativity throughout the folding molecule, with all amides that could be monitored being protected in a concerted fashion. No intermediate partly folded species offering protection to only a proportion of the amides followed was detected at any point during refolding. This is in marked contrast to the previously published results of experiments carried out on unmodified lysozyme (Radford *et al.*, 1992b) and to those carried out here on the unmodified protein under refolding conditions identical to those of $CM^{6,127}$-lysozyme (see Figure 3), in which the hydrogens of the α-domain of the protein are protected more rapidly ($\tau \sim 65$ ms) than those of the β-domain ($\tau \sim 300$ ms). Approximately 20% of the population folds very slowly in the case of both $CM^{6,127}$-lysozyme and the unmodified protein; this has been attributed previously at least in part to peptide bond *cis–trans* isomerization involving the two proline residues in the lysozyme sequence (Kato *et al.*, 1982), but no further analysis of these very slow steps is made here.

A common feature of the protection profiles of the two proteins is their biphasic nature. This suggests that $CM^{6,127}$-lysozyme, like its unmodified counterpart (Radford *et al.*, 1992b; Miranker *et al.*, 1993), folds along alternate, parallel folding pathways. This was investigated further by varying the intensity and the length of the labeling pulse. Pulse intensity experiments were performed at refolding times of 20 and 200 ms by varying the pH of the labeling pulse from pH 9.0 to 10.5. At both refolding times, no significant change in labeling (<5%) was observed over this pH range for the majority of amides monitored, enabling a lower limit of 500 to be put on the protection factors of those amides in molecules that are protected on these time scales. The possibility of an EX1 exchange regime was eliminated by varying the length rather than the pH of the labeling pulse (Roder *et al.*, 1988; Udgaonkar & Baldwin, 1988). Experiments were carried out in which a 16 ms labeling pulse was applied after 23 ms of refolding and a 24 ms pulse after 108 ms of refolding. In both these cases, no significant dependence on labeling time was observed (<5% difference in the extent of labeling from experiments using an 8.4 ms pulse). This indicates that $CM^{6,127}$-lysozyme, like the unmodified protein, folds *via* parallel folding pathways, with the branch point being before any protection from exchange occurs. The Cys6–Cys127 disulfide bridge is not, therefore, the cause of these alternative pathways.

Only three backbone amides were measurably affected by the intensity of the labeling pulse at the 20 ms refolding time point, namely, those of Trp63, Cys64, and Ile78. This suggests that although the fast folding phase affords a high degree of

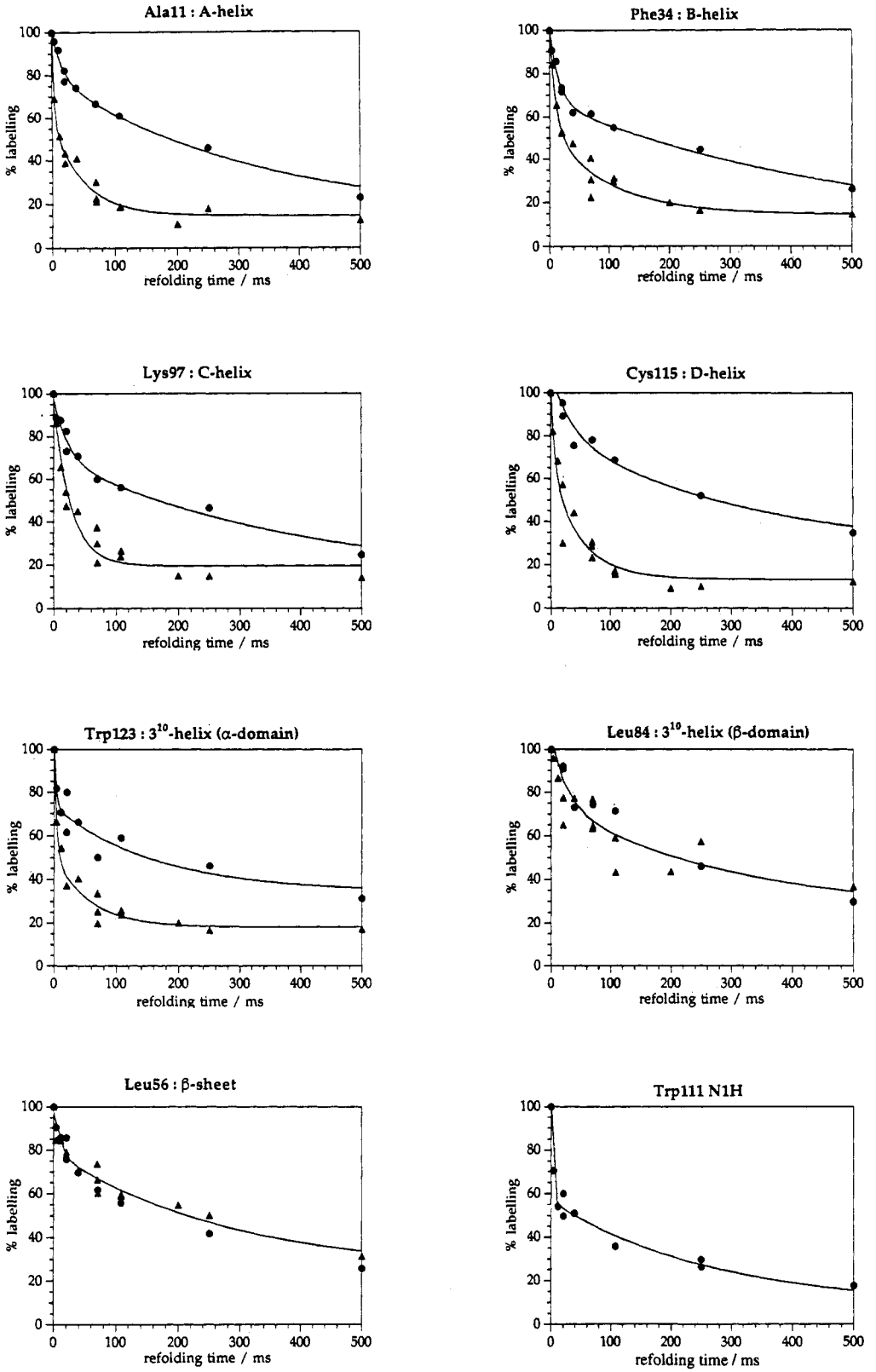


FIGURE 3: Time courses for the protection of representative amides during refolding of $CM^{6,127}$-lysozyme (●) and unmodified hen lysozyme (▲). The extent of labeling is plotted against time after the initiation of refolding. One residue is shown as a representative of each element of secondary structure in the native state. The protection profile for the indole hydrogen of Trp111 (in the α-domain) of the protein is also shown. The curves drawn are the best fits to the sum of two exponentials over the time course studied.

protection to all of the amides in 20–30% of the folding population, these three residues are only partially protected in the species present at this time under these labeling conditions. Behavior of this kind has also been observed for the same residues in the folding of unmodified lysozyme (Radford *et al.*, 1992b) and for the corresponding residues in human lysozyme (Hooke *et al.*, 1994). These three residues are located in the long loop of the native lysozyme structure (residues 61–80), a region constrained by the two disulfide-bridges Cys64–Cys80 and Cys76–Cys94 (Figure 1). Partial protection of these three residues suggests that although protective structure involving these residues forms rapidly, it

is not highly stable against hydrogen exchange under these labeling conditions. Partial protection in this region may be a result of a lack of persistent native-like structure in the vicinity of these disulfide bridges, requiring a rearrangement step to give the fully protected native structure.

The protection of tryptophan indole protons is also a valuable probe of structure formation, particularly since they may be used as indicators of tertiary structure formation in the refolding molecules. Two of the six indole protons (Trp108 and Trp111) of $CM^{6,127}$-lysozyme are sufficiently slow to exchange in the native state to be used as probes for this experiment. In the native state, these indoles are buried within the hydrophobic box region of the protein (Smith *et al.*, 1993). Protection profiles for tryptophan indole protons were measured by integrating the N1H peaks of Trp108 and Trp111 from one-dimensional NMR spectra and normalizing the areas to those of two resolved resonances (Tyr23 $\epsilon$CH and Tyr53 $\epsilon$CH).

The indole proton of Trp108 was found to be protected very rapidly, indeed more rapidly than any of the backbone amides in the protein, but the amplitude of any further change was too small to quantify. Data for Trp111, however, could be fitted to a single-exponential function with a time constant of 300 ± 10 ms, but the amplitude of this phase is only *ca.* 50% of the total change expected (Figure 3), indicating that, within the dead time of the experiment (≤3 ms), approximately half of the population is already sufficiently structured to afford protection to these indole hydrogens. This indicates that although the majority of the molecules do not form stable structure involving backbone amides on this time scale, there clearly is a rapid event, presumably involving substantial hydrophobic collapse, that renders these tryptophan side chains significantly protected from exchange on a very fast time scale. Varying the pH of the labeling pulse from pH 9.0 to 10.5 at 20 ms refolding time gave rise to a change of ~20% in the extent of labeling, indicative of partial protection of these indole protons. It was not possible to calculate protection factors for these indole protons since the pH dependence could not be fitted to a simple mechanism, but an upper limit of 50 can be estimated. Thus, extensive hydrophobic collapse occurs very rapidly, but at these early times the structure around these tryptophan side chains is less protective than that formed at later times for the amide hydrogens.

*Hydrogen Exchange Pulse Labeling Monitored by ESI-MS.* ESI-MS has been developed recently as a means of studying protein folding through hydrogen exchange phenomena that is complementary to NMR (Miranker *et al.*, 1993). Although NMR is a residue-specific technique, it can only reveal proton occupancy at each amide site averaged over all molecules in the sample. Mass spectrometry, by contrast, can distinguish populations of protein molecules in a sample that differ in mass by only a few daltons. This method was particularly important in the folding studies of $CM^{6,127}$-lysozyme since the NMR data show that the $\alpha$- and $\beta$-domains fold to protective structure with identical kinetics and amplitudes in both phases of folding. This behavior is consistent with two fundamentally different folding mechanisms, involving either cooperative folding of the two domains or, alternatively, the $\alpha$- and $\beta$-domains folding independently but with identical kinetics. It is not possible to distinguish between these two possibilities by NMR alone. In contrast, as has been shown for intact lysozyme (Miranker *et al.*, 1993), ESI-MS can distinguish between these two possibilities by the detection of any protected species that are only partially deuterium-labeled.

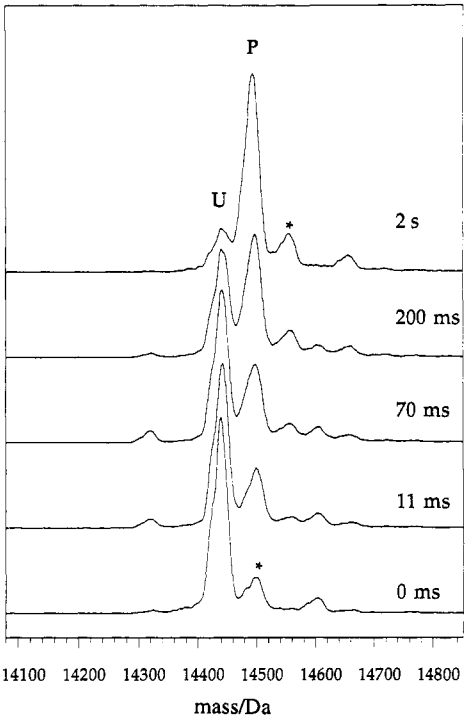


FIGURE 4: Time course for refolding of $CM^{6,127}$-lysozyme as monitored by pulsed amide hydrogen exchange labeling in conjunction with ESI-MS. The peak marked U (14 431 Da) represents the fully unprotected species and that marked P (14 486 Da) the fully protected folded protein. A third peak (*) has been identified, which arises from a species in which a third carboxymethyl group is attached to the side chain of His15 (Yamada *et al.*, 1984). The mass of this species is 14 489 Da in the unprotected state and 14 544 Da in the fully protected state. Peaks of low intensity at higher mass arise from $Na^+$ adducts.

Figure 4 shows mass spectra obtained from samples of $CM^{6,127}$-lysozyme labeled under conditions identical to those in the NMR experiments at various times after the initiation of refolding. At the zero time point, only a single peak is observed at mass 14 431 Da, corresponding to the unprotected $CM^{6,127}$-lysozyme (the small peak at 14 489 Da results from a species carboxymethylated at His15; see Figure 4 legend). The mass of $CM^{6,127}$-lysozyme calculated from the amino acid sequence is 14 423 Da; the difference in mass from that observed for the zero time point arises from the 5% residual deuterium content in the labeling buffer, which is assumed to be distributed randomly throughout the protein. As the refolding time is increased, a second peak emerges with a mass 55 Da greater than that observed for the zero time point. At the same time, the intensity of the peak arising from the fully unprotected protein diminishes. The peak at mass 14 486 Da corresponds to a fully protected species in which 55 slowly exchanging hydrogens, under the refolding conditions employed, become protected from exchange. The mass of this species is heavier than the number of amides (38) that could be individually monitored by NMR, reflecting the problems of spectral overlap and the exchange of amide protons during the acquisition of NMR spectra, which limits the number of amides that could be monitored reliably in the native protein. It is, however, consistent with the number of slowly exchanging amides anticipated under these solution conditions from the data in Figure 2 and may also include some slowly exchanging side chain hydrogens.

The change with time of the relative heights of the two peaks present in the mass spectra is clearly biphasic and is fully consistent with a fast phase with a time constant of ~20

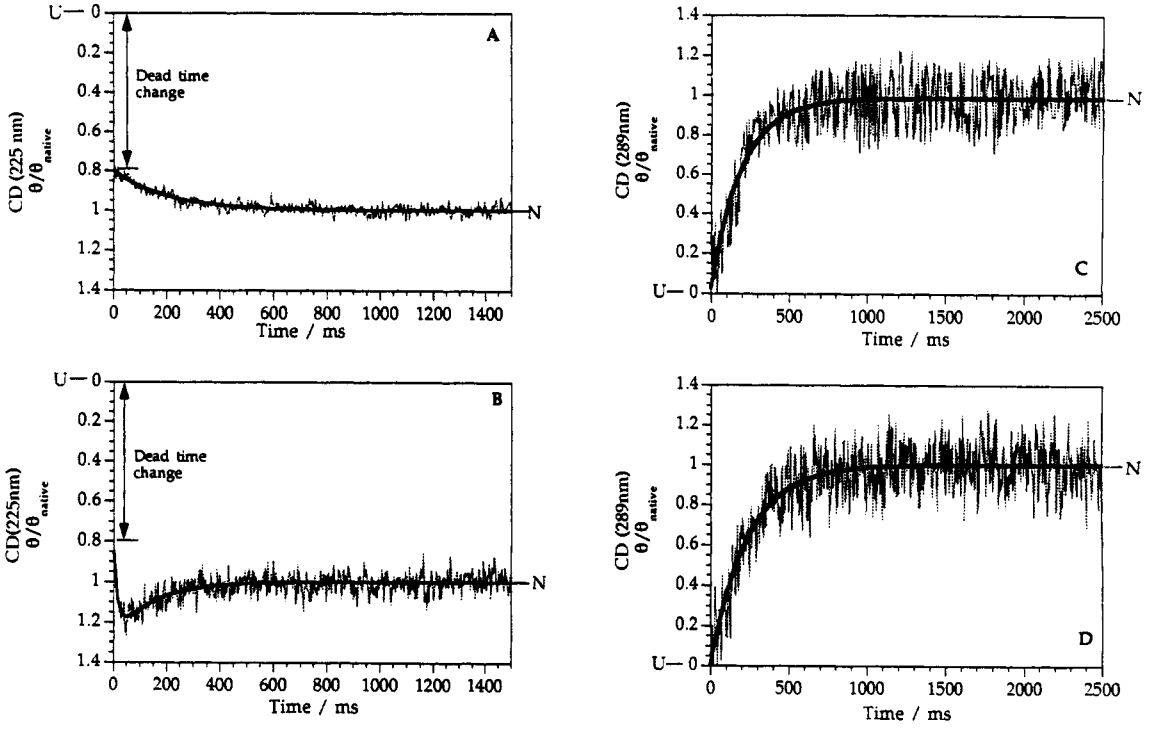


FIGURE 5: Refolding kinetics of $CM^{6,127}$-lysozyme and unmodified lysozyme monitored by stopped-flow CD in the far UV at 225 nm (A and B, respectively) and near UV at 289 nm (C and D, respectively). Data have been normalized to give the fraction of native-like protein present. U represents the ellipticity of the unfolded state and N that of the folded state. The curves represents the best fits to the data.

ms and an amplitude of 20–30%, and a slower phase protecting the remainder of the molecules with a time constant of ~300 ms, in accord with the NMR data presented here. No species of intermediate mass is observed, despite the fact that protection in the $\alpha$- and $\beta$-domains alone should give mass shifts of 28 and 22 Da, respectively. These species would be clearly evident in the mass spectra, provided that they are populated to at least the 5% level (Miranker *et al.*, 1993). This result confirms that the protein indeed folds cooperatively without the formation of any detectable protected transient intermediate states.

*Stopped-Flow Optical Measurements.* Circular dichroism of proteins in the far-UV region arises primarily from secondary structure content and is dominated at 225 nm by $\alpha$-helical structure. The refolding of $CM^{6,127}$-lysozyme monitored at 225 nm indicates the presence of at least two kinetic phases (Figure 5). The CD signal achieves some 80% of its native intensity within the 2 ms dead time of measurement, suggesting that a significant amount of the native helical content in the protein is developed within a very short time after the initiation of refolding. The kinetic trace observed for the remainder of the folding event can be fitted to a single exponential with a time constant of 200 ± 35 ms. These two characteristics are virtually identical to those seen in the refolding of the unmodified protein under identical refolding conditions. These data, however, demonstrate a marked difference between the folding of the two proteins, in that for the unmodified protein a second fast phase is observed ($\tau \sim$ 15 ms) yielding an ellipticity at 225 nm actually greater than that of the native protein. This overshoot previously has been attributed to a contribution to the spectrum from disulfide bridges possibly in a non-native conformation in a partially folded species (Chaffotte *et al.*, 1992). These results, therefore, strongly suggest either that removal of the Cys6–Cys127 disulfide prevents the formation of these interactions or that this disulfide bridge is the major chromophore contributing to the spectrum at this wavelength. The slow phase giving rise to ellipticity characteristic of a native-like state has, however, a remarkably similar time constant for both proteins.

The near-UV signal at 289 nm in the CD spectrum of hen lysozyme in the native state arises primarily from tryptophan residues in fixed orientations (Goux & Hooker, 1980). Kinetic refolding of $CM^{6,127}$-lysozyme monitored at this wavelength adequately fits a single-exponential function with a time constant of 205 ± 20 ms (Figure 5), in good agreement with the slow phase observed by CD at 225 nm. It is interesting to note that although hydrogen exchange protection of tryptophan indoles occurs very rapidly as described earlier, the fact that these do not give rise to a signal in the near-UV CD suggests either that fixed tertiary interactions are not formed at this early stage of folding or that those tryptophans that could be monitored by NMR do not contribute to the CD signal.

The intrinsic tryptophan fluorescence of lysozyme in the native state is thought to arise primarily from Trp62 and Trp108 (Imoto *et al.*, 1971). Figure 6 shows the kinetic trace of total intrinsic tryptophan fluorescence emission during refolding of $CM^{6,127}$-lysozyme. The kinetics exhibit two distinct phases, the first being a dead time change accounting for approximately 80% of the total change in fluorescence intensity. The remaining phase fits well to a single-exponential function, with a time constant of 267 ± 15 ms. This latter phase is similar in rate to that obtained by other optical techniques and to the slowest phase of protection observed by the pulse labeling technique. In contrast, refolding of the unmodified protein under these conditions exhibits three distinct kinetic phases. A significant dead time quenching event, accounting for almost all of the total expected amplitude, is followed by a fast phase ($\tau \sim$ 20 ms) in which fluorescence is quenched beyond that of the native state. Recovery of native state fluorescence intensity then occurs on a 200 ms time scale. These observations are consistent with the rates of folding measured by CD described here and in data published previously (Radford *et al.*, 1992b; Itzhaki *et al.*, 1994) and

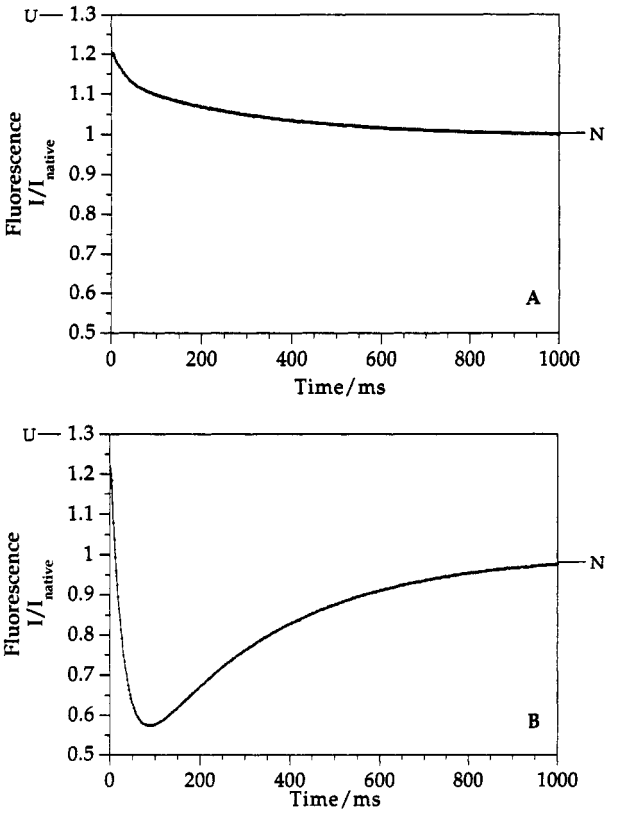


FIGURE 6: Refolding kinetics of $CM^{6,127}$-lysozyme (A) and unmodified lysozyme (B) monitored by stopped-flow intrinsic tryptophan fluorescence. Data have been normalized to the fraction of native state fluorescence, such that the value of the denatured state is 0 and that of the native state is 1.

denaturant concentration (M. E. Denton and H. A. Scheraga, personal communication).

Binding of the fluorescently labeled inhibitor MeU-diNAG to $CM^{6,127}$-lysozyme as a probe for formation of the native protein was also monitored by stopped-flow fluorescence. These data were fitted to a single-exponential function with a time constant of 230 ± 30 ms, in very good agreement with the slow phase observed by intrinsic tryptophan fluorescence and circular dichroism in the near and far UV. No fast phase is observed, however, in this experiment, indicating that the fast phase of amide protection does not give rise to fully native protein, even though both the $\alpha$- and $\beta$-domains are substantially protected from exchange. A similar result has been observed previously in the refolding process of the unmodified protein (Itzhaki *et al.*, 1994).

## DISCUSSION

The folding of lysozyme can be divided into at least three distinct stages, of which we now have some considerable understanding (Dobson *et al.*, 1994). The first steps occur within the dead time of measurement, on a millisecond time scale, and involve at least some degree of hydrophobic collapse, as judged by intrinsic tryptophan fluorescence, fluorescence quenching with iodide, and ANS binding (Itzhaki *et al.*, 1994). The rapid, but partial, protection from exchange of tryptophan indole protons observed in this work is also consistent with such a conclusion. This collapse occurs concomitantly with the formation of a significant amount of secondary structure, probably a near-native complement of $\alpha$-helical structure, as indicated by CD in the far-UV region (Chaffotte *et al.*, 1992; Radford *et al.*, 1992b). No binding of MeU-diNAG inhibitor is observed on this time scale however, indicating, in accord with data in the near-UV CD and the lack of substantial amide hydrogen exchange protection, that few if any fixed tertiary interactions are present in the refolding mixture at this early stage. These states are thought to have similarities to equilibrium partially folded states or molten globules observed for some proteins in which a majority of the secondary structure content is present, yet no fixed tertiary interactions are detectable (Kuwajima, 1992; Baldwin, 1993; Buck *et al.*, 1993). These features are common both to the folding of intact lysozyme and to that of the three-disulfide derivative. Removal of a disulfide bridge in $CM^{6,127}$-lysozyme, therefore, appears to have no significant effect on the very rapid formation of the collapsed state. Interestingly, in this regard, for most of the proteins whose refolding has been studied to date by stopped-flow optical techniques, a collapsed state has been inferred within a few milliseconds of the initiation of folding, despite the fact that the overall folding kinetics of these proteins differs by orders of magnitude (Kuwajima *et al.*, 1993). For example, in lysozyme the majority of molecules appears to fold completely within less than a second, while interleukin-1$\beta$ takes several minutes to acquire its final native structure under the conditions studied (Varley *et al.*, 1993). Nevertheless, in both cases the majority of the secondary structural content of the native protein, at least as judged by far-UV CD, appears to be formed at the earliest possible point of detection, indicating that subsequent folding events are independent of the rate of this initial collapse.

The second observable stage in the folding process of intact lysozyme is the formation of an intermediate species in which the $\alpha$-domain of the protein forms structure protected from exchange in the majority of molecules, whereas the $\beta$-domain is still unstable and not protective toward hydrogen exchange. Protection of the $\alpha$-domain occurs on a time scale similar to that of the development of excess ellipticity in the far-UV CD and of quenching of tryptophan fluorescence intensity to a value greater than that of the native state (Itzhaki *et al.*, 1994). During the refolding of $CM^{6,127}$-lysozyme none of these characteristics is observed. It is difficult to correlate quantitatively the results of the different techniques because, for example, the origin of the chromophores in the far-UV CD and fluorescence is not known, and the different techniques may monitor very different molecular events that occur, by chance, on similar time scales. The fact that none of these characteristics is observed during the folding of $CM^{6,127}$-lysozyme strongly suggests, however, that formation of a stable $\alpha$-domain intermediate is correlated with the development of non-native interactions that quench the fluorescence of tryptophan residues, most of which are located in the $\alpha$-domain, and give rise either directly or indirectly to the overshoot in the ellipticity at 225 nm. It is interesting that this overshoot has been attributed, by reconstruction of the spectrum of the intermediate, to the contribution of disulfide bridge chromophores to the far-UV CD (Chaffotte *et al.*, 1992). It is possible that the lack of such an overshoot in the refolding of $CM^{6,127}$-lysozyme results from the fact that this interaction directly involves the Cys6–Cys127 disulfide bond. It is also possible, however, that the removal of the disulfide bridge destabilizes the $\alpha$-domain intermediate to such an extent that any non-native interactions cannot be detected by these methods.

We have previously shown that removal of the Cys6–Cys127 disulfide bridge has a large effect on the configurational entropy of the native protein, favoring the unfolded state by some 70 ± 5 J $K^{-1}$ $mol^{-1}$ (Cooper *et al.*, 1992). The effects of this are seen in the overall stability of the intact protein and

in the increased rate of hydrogen exchange (see above). It seems extremely likely, therefore, that a similar destabilization will be experienced in the $\alpha$-domain intermediate, in which this disulfide bridge is located. The conditions of the pulse labeling experiment require protection factors of any intermediate species to be greater than about 20 in order to be detectable. Protection factors in the $\alpha$-domain intermediate of intact lysozyme are >500, but the destabilization arising from the removal of Cys6–Cys127 could well reduce these values such that the $\alpha$-domain intermediate cannot be detected in these pulse labeling experiments. In an additional series of experiments, therefore, refolding of $CM^{6,127}$-lysozyme was examined at low temperature and in the absence of denaturant (by dilution from the pH 2 denatured state of $CM^{6,127}$-lysozyme), where residual denaturant cannot be responsible for the destabilization of any intermediate species, in order to attempt to detect protection in an intermediate. Again, however, even under these conditions no evidence for a stable intermediate was observed (S. J. Eyles, S. E. Radford, and C. M. Dobson, unpublished results). The stability of intermediate species in the absence of the Cys6–Cys127 disulfide bridge, therefore, presumably is too low even under these conditions to enable its detection by the experiments employed.

It is interesting to compare these findings with those of recent studies in which the structures of stable disulfide intermediates of $\alpha$-lactalbumin were investigated (Ewbank & Creighton, 1993). $CAM^{6,120}$-bovine $\alpha$-lactalbumin is a three-disulfide carboxyamidomethylated derivative of $\alpha$-lactalbumin lacking the Cys6–Cys120 disulfide bridge homologous to that in $CM^{6,127}$-lysozyme. Structural studies by circular dichroism (Ewbank & Creighton, 1993) and NMR (S. J. Eyles, J. J. Ewbank, T. E. Creighton, S. E. Radford, and C. M. Dobson, unpublished results) indicate that both the native state and the molten globule state of this protein (formed at low pH or in the absence of the calcium ligand) are significantly less stable than their unmodified counterparts. Indeed, the calculated difference in free energy of unfolding of the molten globule states from guanidinium chloride denaturation curves is $\Delta\Delta G(\alpha\text{LA} - \text{CM}^{6,120}\alpha\text{LA}) \approx 5$ kJ mol$^{-1}$ (Ikeguchi & Sugai, 1992), which would result in a significant destabilization of such a marginally stable species.

The third stage in the folding of lysozyme is the formation of the native state of the protein. The overall folding rate of $CM^{6,127}$-lysozyme, remarkably, is quite similar to that of the unmodified protein ($\tau \sim 200$ ms) and is therefore apparently largely unaffected by the removal of the Cys6–Cys127 disulfide bridge. This suggests that the rate-determining step for folding of both proteins does not reflect the stability of the $\alpha$-domain intermediate, but involves stabilization of the $\beta$-domain and the docking of elements of secondary structure to form fixed tertiary contacts in the native folded state. This may in part reflect the fact that $\beta$-structure is intrinsically slower to form than $\alpha$-helical structure (Dill *et al.*, 1993; Varley *et al.*, 1993). In the case of $CM^{6,127}$-lysozyme, the structure of the $\alpha$-domain is sufficiently labile in the absence of the cross-linking disulfide bridge that it can only be observed concomitantly with stabilization of the $\beta$-domain. The fact, however, that some lysozyme molecules are able to protect amides rapidly in the fast phase of folding suggests that misfolding events may play an important role and that only some topologies formed in this initial ensemble of collapsed states can proceed directly to the fully protected state. It is interesting in this regard that it has been suggested that the most highly protected amides in the native state of a protein represent the earliest structural region to form during refolding (Woodward, 1994). This theory would predict, therefore, in the light of the hydrogen exchange data given here, that the B- and C-helices and the third strand of the main $\beta$-sheet in $CM^{6,127}$-lysozyme should be the earliest regions to fold. The results described in this paper are not, however, consistent with this prediction; indeed, residues Trp63, Cys64, and Ile78 are highly protected from exchange in the native protein, but are in fact the last amides to attain full protection from exchange during folding. If this region is the earliest structure to form during folding, therefore, it cannot be of sufficient persistent stability to protect against hydrogen exchange. This result could be a consequence of the rate-limiting step in folding arising from misfolding events, rather than intrinsic features of the native structure. Whatever the explanation, this result reaffirms the independent nature of the two folding domains in lysozyme. Further characterization of the effects of specific residues in mutants and naturally occurring variant lysozymes will allow these individual submolecular kinetic events to be examined more fully (S. D. Hooke, S. J. Eyles, A. Miranker, C. V. Robinson, and C. M. Dobson, manuscript in preparation).

One of the particularly interesting features in the study of protein folding is relating *in vitro* studies to the processes occurring within the living cell. Surprisingly little is known, at a molecular level, about how a protein attains its final native structure as it emerges from the ribosome and, for disulfide-containing proteins, is transferred into the lumen of the endoplasmic reticulum. A number of helper proteins, including protein-disulfide isomerase, play an important part in folding and processing in this protein-rich cell compartment (Freedman, 1992; Jaenicke, 1993). As far as disulfide formation in the ER lumen is concerned, it is evident that the two cysteine residues must come into close proximity in a partly folded state in order to allow the oxidative formation of a disulfide bridge. As regions of the protein begin to fold, disulfides can form concomitantly, driving the folding process toward completion by stabilizing local structure relative to the unfolded state. An important, perhaps dominant, contribution to the driving force is the greater reduction in entropy in more unfolded states that results from a cross-linking reaction.

In the case of hen lysozyme, extensive studies have suggested that the Cys6–Cys127 disulfide is the last to form, at least in the case of *in vitro* oxidative refolding experiments (Anderson & Wetlaufer, 1976). In this paper, we have shown that the presence of this cross-link is not required for the complete folding of the protein since the three-disulfide species clearly folds efficiently to its native state. Instead, its formation creates a constraining loop that confers stability on the final folded conformation and on partially folded intermediates by entropically disfavoring the unfolded state (Cooper *et al.*, 1992). The role of the disulfide bridge in this case appears to be one of maintenance of folded structure rather than inducing structure into the polypeptide chain. Further folding studies of proteins both in the presence and absence of disulfide bonds should determine whether this is a general role of disulfide bridges in the folding of proteins or whether disulfide bonds in different environments have very different roles.

## REFERENCES

Anderson, W. L., & Wetlaufer, D. B. (1976) *J. Biol. Chem. 251*, 3147–3153.

Aue, W. P., Bartholdi, F., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75–86.

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

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.
Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, R. (1965) *Nature 206*, 757–761.
Buck, M., Radford, S. E., & Dobson, C. M. (1993) *Biochemistry 32*, 669–678.
Chaffotte, A. F., Guillou, Y., & Goldberg, M. E. (1992) *Biochemistry 31*, 9694–9702.
Clarke, J., & Fersht, A. R. (1993) *Biochemistry 32*, 4322–4329.
Cooper, A., Eyles, S. J., Radford, S. E., & Dobson, C. M. (1992) *J. Mol. Biol. 225*, 939–943.
Creighton, T. E. (1992) *Science 256*, 111–112.
Denton, M. E., & Scheraga, H. A. (1991) *J. Protein Chem. 10*, 213–232.
Dill, K. A., Fiebig, K. M., & Chan, H. S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1942–1946.
Dobson, C. M., Evans, P. A., & Radford, S. E. (1994) *Trends Biochem. Sci. 19*, 31–37.
Elöve, G. A., Chaffotte, A. F., Roder, H., & Goldberg, M. E. (1992) *Biochemistry 31*, 6876–6883.
Evans, P. A., & Radford, S. E. (1994) *Curr. Opin. Struct. Biol. 4*, 100–106.
Ewbank, J. J., & Creighton, T. E. (1993) *Biochemistry 32*, 3694–3707.
Fan, P., Bracken, C., & Baum, J. (1993) *Biochemistry 32*, 1573–1582.
Fersht, A. R. (1993) *FEBS Lett. 325*, 5–16.
Fersht, A. R., & Serrano, L. (1993) *Curr. Opin. Struct. Biol. 3*, 75–83.
Freedman, R. B. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 457–541, W. H. Freeman, New York.
Goux, W. J., & Hooker, T. M., Jr. (1980) *Biopolymers 19*, 2191–2208.
Harding, M. M., Williams, D. H., & Woolfson, D. N. (1991) *Biochemistry 30*, 3120–3128.
Heinrikson, R. L., & Meredith, S. C. (1984) *Anal. Biochem. 136*, 65–74.
Hill, C. P., Johnston, N. L., & Cohen, R. E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4136–4140.
Hooke, S. D., Radford, S. E., & Dobson, C. M. (1994) *Biochemistry 33*, 5867–5876.
Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544–1548.
Ikeguchi, M., & Sugai, S. (1992) *Biochemistry 31*, 12695–12700.
Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6965–6972.
Imoto, T., & Rupley, J. A. (1973) *J. Mol. Biol. 80*, 657–667.
Imoto, T., Forster, L. S., Rupley, J. A., & Tanaka, F. (1971) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1151–1155.
Itzhaki, L. S., Evans, P. A., Dobson, C. M., & Radford, S. E. (1994) *Biochemistry 33*, 5212–5220.
Jaenicke, R. (1993) *Curr. Opin. Struct. Biol. 3*, 104–112.
Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892–895.
Kato, S., Okamura, M., Shimamoto, N., & Utiyama, H. (1981) *Biochemistry 20*, 1080–1085.
Kato, S., Shimamoto, N., & Utiyama, H. (1982) *Biochemistry 21*, 38–43.
Khorasanizadeh, S., Peters, I. D., Butt, T. R., & Roder, H. (1993) *Biochemistry 32*, 7054–7063.
Kim, K.-S., & Woodward, C. (1993) *Biochemistry 32*, 9609–9613.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Kuwajima, K. (1992) *Curr. Opin. Biotechnol. 3*, 462–467.
Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.
Kuwajima, K., Semisotnov, G. V., Finkelstein, A. V., Sugai, S., & Ptitsyn, O. B. (1993) *FEBS Lett. 334*, 265–268.
Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.
Matthews, B. W. (1993) *Curr. Opin. Struct. Biol. 3*, 589–593.
Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., & Dobson, C. M. (1993) *Science 262*, 896–899.
Moore, P. B., Traut, R. R., Noller, H., Pearson, P., & Delius, H. (1968) *J. Mol. Biol. 31*, 441–461.
Pedersen, T. G., Sigurskjold, B. W., Andersen, K. V., Kjaer, M., Poulsen, F. M., Dobson, C. M., & Redfield, C. (1991) *J. Mol. Biol. 218*, 413–426.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20–24.
Radford, S. E., Woolfson, D. N., Martin, S. R., Lowe, G., & Dobson, C. M. (1991) *Biochem. J. 273*, 211–217.
Radford, S. E., Buck, M., Topping, K. D., Evans, P. A., & Dobson, C. M. (1992a) *Proteins: Struct., Funct., Genet. 14*, 237–248.
Radford, S. E., Dobson, C. M., & Evans, P. A. (1992b) *Nature 358*, 302–307.
Redfield, C., & Dobson, C. M. (1988) *Biochemistry 27*, 122–136.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700–704.
Smith, L., Sutcliffe, M. J., Redfield, C., & Dobson, C. M. (1993) *J. Mol. Biol. 229*, 930–944.
Strausberg, S., Alexander, P., Wang, L., Gallagher, T., Gilliland, G., & Bryan, P. (1993) *Biochemistry 32*, 10371–10377.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694–699.
Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., & Clore, G. M. (1993) *Science 260*, 1110–1113.
Wedin, R. E., Delepierre, M., Dobson, C. M., & Poulsen, F. M. (1982) *Biochemistry 21*, 1098–1103.
Weissman, J. S., & Kim, P. S. (1992) *Science 256*, 112–114.
Woodward, C. K. (1994) *Curr. Opin. Struct. Biol. 4*, 112–116.
Yamada, H., Uozumi, F., Ishikawa, A., & Imoto, T. (1984) *J. Biochem. 95*, 503–510.