Protein Science (1993), *2*, 869–876. Cambridge University Press. Printed in the USA.
Copyright © 1993 The Protein Society

STEIN AND MOORE AWARD ADDRESS

The molten globule intermediate of apomyoglobin and the process of protein folding

DOUG BARRICK AND ROBERT L. BALDWIN
Department of Biochemistry, Beckman Center, Stanford University, Stanford, California 94305-5307
(RECEIVED January 14, 1993; REVISED MANUSCRIPT RECEIVED February 23, 1993)

Abstract

The molten globule model for the beginning of the folding process, which originated with Kuwajima's studies of α-lactalbumin (Kuwajima, K., 1989, *Proteins Struct. Funct. Genet. 6*, 87–103, and references therein), states that, for those proteins that exhibit equilibrium molten globule intermediates, the molten globule is a major kinetic intermediate near the start of the folding pathway. Pulsed hydrogen-deuterium exchange measurements confirm this model for apomyoglobin (Jennings, P.A. & Wright, P.E., in prep.). The energetics of the acid-induced unfolding transition, which have been determined by fitting a minimal three-state model (N ⇔ I ⇔ U; N = native, I = molten globule intermediate, U = unfolded) show that I is more stable than U at neutral pH (Barrick, D. & Baldwin, R.L., 1993, *Biochemistry 32*, in press), which provides an explanation for why I is formed from U at the start of folding. Hydrogen exchange rates measured by two-dimensional NMR for individual peptide NH protons, taken together with the CD spectrum of I, indicate that moderately stable helices are present in I at the locations of the A, G, and H helices of native myoglobin (Hughson, F.M., Wright, P.E., & Baldwin, R.L., 1990, *Science 249*, 1544–1548). Directed mutagenesis experiments indicate that the interactions between the A, G, and H helices in I are loose (Hughson, F.M., Barrick, D., & Baldwin, R.L., 1991, *Biochemistry 30*, 4113–4118), which can explain why I is formed rapidly from U at the start of folding. These experiments are consistent with the explanation proposed earlier (Baldwin, R.L., 1989, *Trends Biochem. Sci. 14*, 291–294) for the stabilization of native secondary structures in molten globule intermediates, namely that each unit of native secondary structure has a hydrophobic face, and the hydrophobic surfaces of two units of secondary structure can interact loosely to provide mutual stabilization. These topics are discussed here in the light of some additional results.

Keywords: apomyoglobin; energetics of folding; folding pathways; molten globule; protein folding

It is now generally agreed that structured intermediates are observable during the kinetics of folding of most small proteins. If they are on-pathway intermediates, then it is possible to determine the folding pathway by characterizing their structures. The method of pulsed H/D exchange, which measures the exchange rates of individual peptide NH protons in a folding intermediate using the 2D ^{1}H-NMR spectrum taken after folding is complete, gives detailed structural information about folding intermediates. We argue here, in agreement with the analysis by Kuwajima (1989, and references therein), that equilibrium molten globule intermediates are probably early intermediates in the kinetic process of folding (see also Jennings & Wright [1992 and in prep.] and Baldwin [1993]). Molten globule intermediates have compact structures and substantial amounts of secondary structure, but little if any fixed tertiary structure (Ptitsyn, 1987; Kuwajima, 1989). Because the structure and behavior of an equilibrium intermediate can be determined accurately, we focus attention on equilibrium molten globule intermediates. By studying them, it is possible to connect the results with kinetic studies of the folding process and obtain in-depth information about the initiation of folding.

Reprint requests to: Robert L. Baldwin, Department of Biochemistry, Beckman Center, Stanford University, Stanford, California 94305-5307.

Abbreviations: Mb, myoglobin; apoMb, apomyoglobin; holoMb, holomyoglobin; 2D, two-dimensional; H-bond, hydrogen bond; H/D, hydrogen/deuterium; ANS, 1-anilinonaphthalene-8-sulfonic acid; α-LA, α-lactalbumin; cyt *c*, cytochrome *c*; pH_m, pH midpoint of an unfolding transition; pH*, pH measured in ^2H_2O without correction for the isotope effect; GdmCl, guanidinium chloride.

Structural information about kinetic folding intermediates

Studies of the kinetic folding pathways of seven proteins, monitored by pulsed H/D exchange, have now been reported from six laboratories and studies of more proteins are in progress (Baldwin, 1993). Representative data are shown in Figure 1A for barnase (Bycroft et al., 1990) and in Figure 1B for T4 lysozyme (Lu & Dahlquist, 1992). The kinetics of folding are monitored individually for various NH protons in the polypeptide backbone. Resistance to exchange with solvent is determined at each time point. The exchange rate of a NH proton in the native species depends chiefly on whether it is H-bonded in secondary structure, either in an α-helix or β-sheet, and on the extent of burial of the NH proton inside the protein, away from solvent (see reviews: Woodward et al., 1982; Englander & Kallenbach, 1984). These factors are expected to determine the exchange rates of NH protons in folding intermediates.

Figure 1 provides examples of structural information obtained from pulsed H/D exchange studies of folding intermediates. In Figure 1A, individual peptide NH protons of barnase are seen to become protected against H/D exchange more rapidly during folding if in native barnase they form H-bonds within a unit of secondary structure than if they form tertiary H-bonds. In Figure 1B, T4 lysozyme forms a folding intermediate very rapidly (within the stopped-flow mixing time) that shows only moderate protection against H/D exchange, and various NH protons undergo exchange at different rates, depending on the stability of the surrounding structure.

Our attention was drawn to the molten globule model for early folding intermediates through analyzing the energetics of a molten globule intermediate of apoMb (Barrick & Baldwin, 1993). The molten globule model for early transient intermediates was put forward by Kuwajima (1989 and references therein), and it has been supported by recent work of Jennings and Wright (1992 and in prep.) who find the same protected NH protons in an early folding intermediate of apoMb as in the equilibrium molten globule species. Ptitsyn et al. (1990) proposed that equilibrium molten globule species are not among the earliest intermediates on the kinetic pathway of folding but rather occur as later intermediates. Their proposal assumed that the fluorescent dye ANS binds only to molten globule species; they observed that ANS binding occurs at a relatively late stage in the folding reactions of several proteins. Further study showed, however, that ANS binding also occurs in a burst phase, within the stopped-flow mixing time (a few milliseconds) (Semisotnov et al., 1991), as well as in a later stage of folding. Thus, ANS binding occurs both to early and later folding intermediates.

Molten globule intermediates

Bovine α-LA has long been the paradigm for studies of equilibrium folding intermediates (Kuwajima et al., 1976; Kuwajima, 1977, 1989 and references therein; Dolgikh et al., 1981; Ptitsyn, 1987). The equilibrium intermediate or acid form (A-form) is the sole species present at acid pH, at various salt concentrations, and a similar form can be observed at neutral pH, as an equilibrium intermediate in the GdmCl-induced unfolding of native α-LA (Kuwajima, 1977). Kuwajima (1977) recognized that the A-form is probably similar to the early intermediate seen in kinetic folding experiments at neutral pH, and this proposal was strengthened by the work of Ikeguchi et al. (1986a), who showed that the stability to GdmCl-induced unfolding of the early kinetic intermediate is the same as that of the equilibrium intermediate. The structure of the

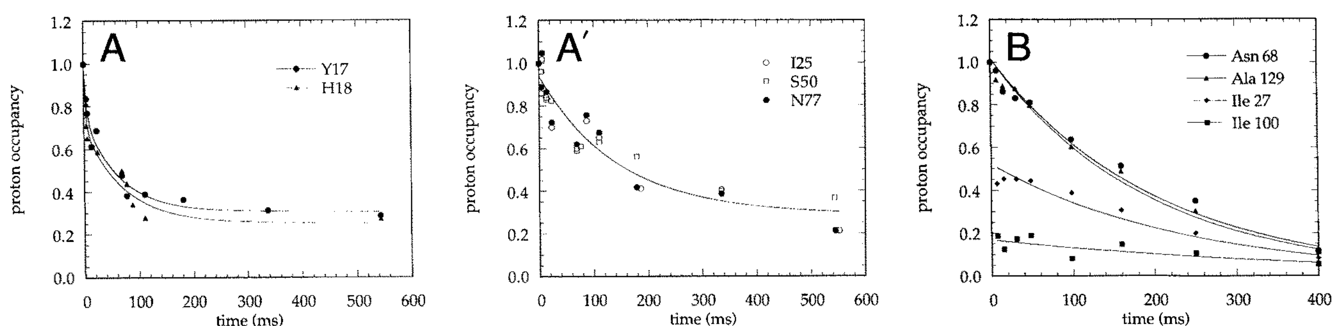

Fig. 1. Observation of kinetic folding intermediates by monitoring the exchange (^2H–^1H) of individual peptide NH protons (present as the ^2H isotopic form in the unfolded protein) following dilution from ^2H_2O to ^1H_2O at various times after the start of folding. **A, A′**: Folding of barnase (Bycroft et al., 1990, redrawn) at 1.3 M urea, pH* 6.3, 25 °C. The peptide NH groups in Tyr 17 and His 18 make H-bonds within an α-helix; those in Ile 25, Ser 50, and Asn 77 make tertiary H-bonds. **B:** Folding of T4 lysozyme (Lu & Dahlquist, 1992, redrawn) in 1.5 M urea, 20 °C, pH 5.1. At various times after the start of folding, $^2H \rightarrow {}^1H$ exchange was initiated by raising the pH to 9.0 (exchange is base catalyzed). The folding intermediate is formed within the stopped-flow mixing time, and the different extents of exchange reflect the different extents to which these NH protons are protected against exchange in the folding intermediate. Ile 27 is in the N-terminal β-sheet, Ile 100 is in helix E, Asn 68 is in helix C, and Ala 129 is in helix G of T4 lysozyme.

A-form of guinea pig α-LA is being analyzed by NMR by Baum, Dobson, and coworkers (Baum et al., 1989; Chyan et al., 1993). Ptitsyn and coworkers pointed out that the A-form of α-LA belongs to a general class of equilibrium folding intermediates that have compact conformations, high contents of secondary structure measured by CD, and little if any fixed tertiary structure (Dolgikh et al., 1981; Ptitsyn, 1987). Ohgushi and Wada (1983) gave the name "molten globule" to this class of folding intermediates when they studied the properties of the acid molten globule intermediate of horse cyt *c*.

Later, CD spectra and NH exchange rates measured by 2D NMR showed that molten globule intermediates of some proteins, including guinea pig α-LA, contain stable α-helices at the locations where helices are found in the native protein. This observation strongly suggests that these "structured molten globules" are plausible intermediates on the kinetic pathway of folding. Stable NH protons are found at the locations of native helices in guinea pig α-LA (Baum et al., 1989; Chyan et al., 1993), cyt *c* (Jeng et al., 1990), and sperm whale apoMb (Hughson et al., 1990).

Structure of the apomyoglobin intermediate

NH exchange results for the molten globule intermediate (I) of apoMb indicate that helices corresponding to the A, G, and H helices of Mb are present but that segments corresponding to the B and E helices are unstructured (Hughson et al., 1990). Because the A, G, and H helices form a compact subdomain in Mb, the results at first suggested that this subdomain is stable when the remainder of apoMb unfolds. Directed mutagenesis was used (Hughson et al., 1991) to find out if the intermediate is stabilized via close-packing interactions between helices, when side chains of the three helices are interdigitated with each other at the contact sites seen in the crystal structure of holoMb (Kinemage 1; Takano, 1977). An A·H helix contact site in native Mb was chosen for study that satisfies the "ridge-into-groove" helix pairing model of Chothia et al. (1981): a spiral ridge of side chains from the H helix fits into a matching groove in the A helix. Ala 130, which occurs near the center of the site, was chosen for study (see Kinemage 2). The G and H helices of Mb lie nearly parallel to each other, connected by a corner loop, with contact sites involving more than four helical turns. Residues Ser 108 and Phe 123 within these G·H contact sites were also chosen for study. Phe 123 is a highly conserved residue in the globin family, whose side chain is buried in the corner loop connecting the G and H helices (Kinemage 3); Ser 108 is near the center of the G·H interface (Kinemage 4). The mutations were chosen to test the effects of varying side-chain size and polarity in disrupting contacts between helices.

The urea and acid unfolding curves for native apoMb were measured first to determine if the mutations destabilize the native structure as expected. Parallel results were obtained by measuring either the urea unfolding curve at neutral pH, which does not resolve I, or the acid unfolding curve, which resolves the N ⇔ I and I ⇔ U transitions. Introducing either a large nonpolar side chain (Trp) or a small one (Gly) in place of Phe 123 strongly destabilizes N. Introducing a charged side chain, such as Lys^+, in place either of Ala 130 or Phe 123 strongly destabilizes N. The results obtained with mutations at Ser 108 gave anomalously small changes in the stability of N, and sometimes its stability was increased. There is a cavity next to Ser 108 in the Mb structure, and the anomalous results were probably caused by alternative close-packing arrangements within this cavity.

The effects of the mutations on the stability of I were surprising: increasing the size of a nonpolar side chain did not destabilize I. The effects of the mutations were monitored by changes in the pH_m of the I ⇔ U transition. Small increases in the stability of I relative to U were usually found when a larger nonpolar side chain was introduced (Fig. 2). Introducing a charged side chain was found to destabilize I relative to U, but the change in the pH_m of the I ⇔ U transition was always smaller than the change in pH_m of the N ⇔ I transition (Fig. 2). We concluded that the A, G, and H helices in I are stabilized by some mechanism other than the close-packing interactions between helices seen in the structure of native myoglobin (Hughson et al., 1991). Although surprising, this conclusion fits in with evidence from other studies of molten globule intermediates that indicates that fixed tertiary structure is absent (Ptitsyn, 1987; Kuwajima, 1989). Ewbank and Creighton (1991) reported that breaking a specific disulfide bond in α-LA produces a molten globule species at neutral pH and more than one disulfide pair-

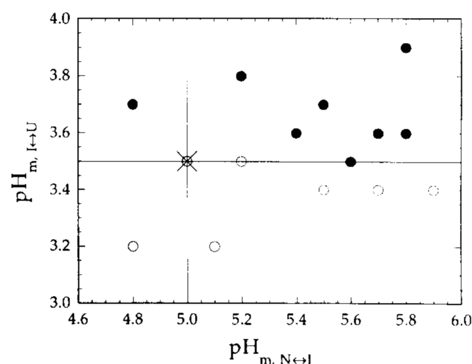

Fig. 2. The pH midpoint of the I ⇔ U transition of apoMb is plotted against the pH midpoint of the N ⇔ I transition for a series of mutations at contact sites between helices in native Mb (data from Hughson et al., 1991: see their Table 1). The figure contrasts the effect of increasing the size of a nonpolar side chain (○) (destabilizing for N ⇔ I but stabilizing for I ⇔ U) with substituting a more polar side chain (●) (destabilizing both for N ⇔ I and I ⇔ U). The data point for wild-type apoMb is marked by ×. An increase in pH_m above the value shown by wild-type protein represents a decrease in stability.

ing is compatible with the molten globule conformation, as expected if the molten globule is stabilized in some way other than by fixed tertiary interactions.

It should be pointed out that the folding intermediate of apoMb is unusual in two respects, when compared to other molten globule intermediates. First, it unfolds when the pH is lowered below 3, and it represents the only case so far of a molten globule intermediate that can be caused to unfold by exposing it to low pH. Second, I is stable at low salt concentrations (2 mM Na citrate is used in the experiments discussed here), whereas typical molten globule intermediates occur at high salt concentrations, where they are stabilized by anion binding (Goto et al., 1990). The acid-unfolded form of horse apoMb at pH 2 does become a molten globule species when 0.5 M KCl is added (Goto & Fink, 1991). The fact that the I form of apoMb is stable at low salt concentrations at pH 4 means that it is an unusually stable molten globule species.

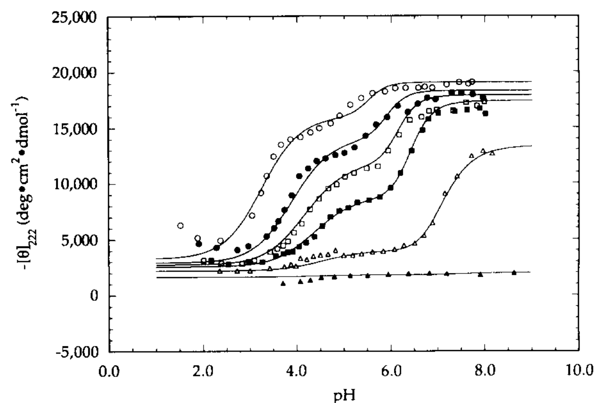

Fig. 3. Acid-induced unfolding transitions curves for sperm whale apoMb at various urea concentrations, 0 °C, 2 mM Na citrate (Barrick & Baldwin, 1993). From left (○) to right (△), the transition curves show unfolding in 0.0, 1.0, 1.5, 2.0, and 3.0 M urea, respectively; the baseline curve for complete unfolding (▲) is measured in 4.5 M urea. The molten globule intermediate I is most populated near pH 4.5 at 0 M urea but is populated between pH 5 and pH 6 in 3 M urea. The solid curves were obtained by fitting the data to a minimal three-state model (N ⇔ I ⇔ U), and the fitted model gives the relative concentrations of N, I, and U in the range of conditions shown here.

Energetics of the apomyoglobin folding intermediate

To test the conclusion that close-packing interactions between helices are lost in the N ⇔ I transition of apoMb, and to measure quantitatively the stability of I in refolding conditions, we undertook a study of the energetics of the N ⇔ I ⇔ U unfolding transition of apoMb (Barrick & Baldwin, 1993). Acid unfolding curves were measured at a series of urea concentrations, and urea unfolding curves were measured at a series of pH values. The results were fitted globally to a minimal three-state model (N ⇔ I ⇔ U). The model has as its unknown parameters the pK_a values and number of titrating groups responsible for the N ⇔ I and I ⇔ U unfolding reactions, plus two parameters for each unfolding reaction, ΔG^0 (0) and m, that describe the linear dependence on urea concentration of the difference between N and I, or I and U, in Gibbs free energy ΔG^0.

The acid unfolding transitions measured at different urea concentrations are shown in Figure 3, together with the curves obtained by fitting the data to a three-state model. The agreement between the fitted curves and the data is good over a wide range of pH and urea concentrations. Thus, the fitted three-state model can be used to explore the behavior of I in various conditions. A basic question is whether or not the molten globule conformation is stable relative to U at neutral pH. Equilibrium molten globule species are typically populated at low pH, because N must become unstable before the molten globule species can be populated. Figure 3 shows that I is populated at progressively higher pH values as the urea concentration is increased; in 3 M urea I is populated at pH 6. Fitting the data to the three-state model shows that I is definitely more stable than U at neutral pH: ΔG^0 for the formation of I from U is -2.1 kcal/mol at 0 M urea, pH 8.0, 0 °C.

Because we have shown that I is more stable than U in refolding conditions, and I contains helical structure found in the native protein, I is expected to be an early intermediate on the folding pathway from U to N, provided I is formed rapidly from U. This expectation is confirmed by the work of Jennings and Wright (1992 and in prep.), who measured the kinetics of stopped-flow refolding of apoMb at pH 6.1, using both pulsed H/D exchange and CD at 222 nm. A structured intermediate is formed within the stopped-flow mixing time: it has protected NH protons within the A, G, and H helices of Mb, but not within the E helix, and only a few NH protons in the B helix show protection. At later times more protected NH protons are found first within the B helix of Mb and then within the E helix (P.E. Wright, pers. comm.). These results indicate that the equilibrium molten globule intermediate of apoMb is formed within the burst phase of stopped-flow refolding, and afterward folding follows a sequential mechanism.

The results for apoMb are parallel to early results of Kuwajima and coworkers (Kuwajima, 1989, and references therein) for the molten globule intermediate of α-LA. They found that the A form is stable relative to the unfolded form D at neutral pH: ΔG^0 for the formation of A from D is -1.4 kcal/mol at pH 7.0, 25 °C, 0.1 M Na^+ (Ikeguchi et al., 1986b). Moreover, they later were able to measure the stability to unfolding by GdmCl of the early kinetic folding intermediate (Ikeguchi et al., 1986a), which equilibrates with the unfolded form before the next step in folding takes place. The early kinetic intermediate has similar spectral properties measured by CD and tryptophan absorbance to the equilibrium molten

globule intermediate and both intermediates have the same stability to GdmCl-induced unfolding (Fig. 4).

Results from David Shortle's laboratory add to the view that unfolded proteins form compact intermediates in physiological folding conditions. Truncated deletion mutants of staphylococcal nuclease are unable to form a native-like structure unless specific stabilizing ligands are added; nevertheless they assume compact, partially structured conformations (Flanagan et al., 1992). Studies of point mutants (Shortle & Meeker, 1989) show how specific side chains influence this compact conformation.

To sum up, because molten globule intermediates are stable relative to the corresponding unfolded forms at neutral pH, and because the absence of fixed tertiary interactions makes it possible to form these intermediates rapidly, molten globules are probable early intermediates in the kinetic process of folding. If molten globule species are more stable than their unfolded forms in folding conditions, then the commonly used two-state model (U ⇔ N) of folding is only an approximation to a more generally applicable three-state model (U ⇔ I ⇔ N) as Kuwajima (1989) pointed out. An important reason for studying the structures and behavior of equilibrium molten globule intermediates is that they do not change with time. It is likely that, in some folding reactions, a series of transient intermediates are populated during the kinetic process of folding (see Elöve et al., 1992; Radford et al., 1992; review, Baldwin, 1993). When this is the case, it is difficult to isolate and study a particular kinetic intermediate.

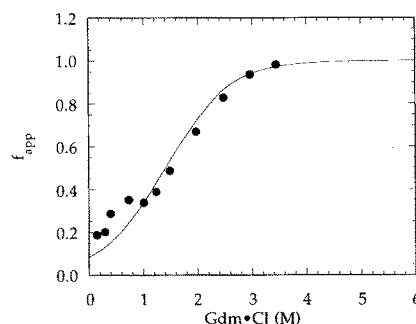

Fig. 4. GdmCl-induced unfolding transition of the kinetic folding intermediate (●) of bovine α-LA compared with that of the equilibrium folding intermediate (solid line) (Ikeguchi et al., 1986a, redrawn). Conditions: 4.5 °C, pH 7.0, 0.05 M NaCl, 0.05 M Na cacodylate. The refolding process at various concentrations of GdmCl is monitored by CD at 222 nm.

Interactions stabilizing the apomyoglobin molten globule intermediate

The conclusion that close-packing interactions between helices are lost when the molten globule intermediate of apoMb is formed is based on the assumption that the pH midpoint of the I ⇔ U unfolding transition responds sensitively to changes in the stability of I. To test this assumption, the fitted three-state model was used to generate a family of curves that show that the effect of destabilizing only N relative to I (Fig. 5A) or both N and I relative to U (Fig. 5B) by successive increments in $-\Delta G^0$ of 0.5 kcal/mol. These curves show that a destabilizing mutation affecting the stability of N relative to I changes only the pH_m of the N ⇔ I transition; on the other hand, although a mutation affecting I relative to U does change the pH_m of the I ⇔ U transition, it may also affect the shape of the entire N ⇔ I ⇔ U unfolding transition.

The unfolding curves for some of the mutants at Ala 130, Ser 108, and Phe 123 (Hughson et al., 1991) are shown in Figure 6. They have the same shape as the predicted curves in Figure 5A for mutations that affect only the N ⇔ I reaction, and they are quite different from the predicted curves in Figure 5B, for mutations that affect only the I ⇔ U reaction. Thus, the fitted three-state model for the energetics of the N ⇔ I ⇔ U unfolding transition supports the conclusion that close-packing interactions are lost in the N ⇔ I transition. Therefore, some other interaction involving the A, G, and H helices must stabilize them in the molten globule intermediate.

The problem of finding the interactions that stabilize the molten globule intermediate of apoMb has been studied by molecular dynamics simulations, both by Tirado-Rives and Jorgensen (1993) and by Brooks (1992). Tirado-Rives and Jorgensen simulated the unfolding kinetics of native apoMb and found a populated intermediate that was similar to the molten globule intermediate. Brooks simulated the dynamics of native apoMb and found that the A, G, and H helices comprise a particularly stable subdomain.

Peptides with sequences corresponding to the individual A, G, and H helices do not form stable helices in isolation. The wild-type H sequence shows about 25% helix formation at 0 °C (Waltho et al., 1989; Hughson et al., 1991) but the wild-type A and G sequences tend to aggregate. An upper limit on the extent of helix formation by the A and G peptides can be determined from peptides with variant sequences. The substitutions in Table 1 are chosen both to enhance helix formation and to reduce aggregation. These experiments show that the wild-type G sequence is a very weak helix former, as reported earlier by Hughson et al. (1991), and the wild-type A sequence is even weaker: its CD spectrum is unlike those shown by partly helical peptides. Therefore, some interaction between the A, G, and H helices must stabilize them in the molten globule species. As argued earlier (Hughson et al., 1991), a likely stabilizing interaction is a loose hydrophobic interaction between the hydrophobic faces of these helices. Mutations that introduce a charged side chain tend to destabilize I, whereas mutations that introduce a large hydrophobic side chain tend to stabilize I (Hughson et al., 1991).

These experiments support the following generalization (Baldwin, 1989) about the mechanism of protein folding:

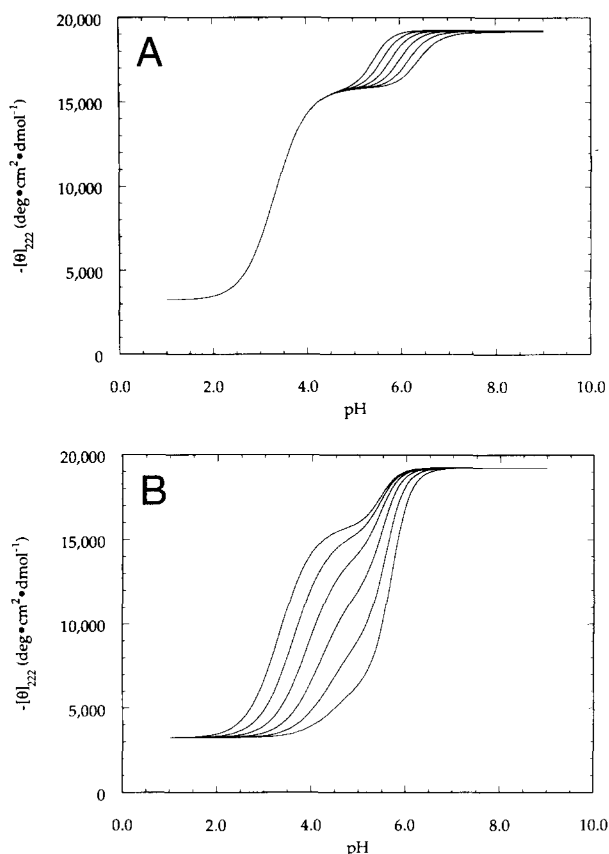

Fig. 5. A family of curves showing the expected effect of a destabilizing mutation in apoMb that affects either the N ⇔ I reaction (**A**) or the I ⇔ U reaction (**B**). Successive curves show the effect of decreasing ΔG^0 for either reaction by 0.5 kcal/mol. These curves are generated by the three-state model used to fit the data in Figure 3. In each case, the most stable protein is represented by the curve on the far left and the least stable one by the curve on the far right.

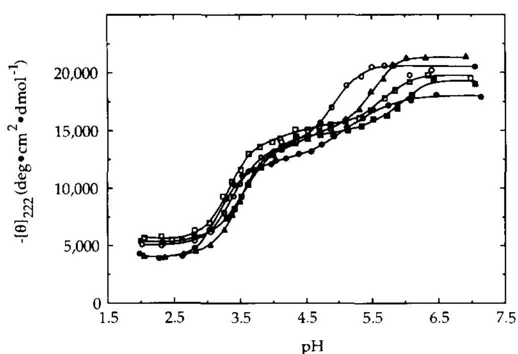

Fig. 6. Acid unfolding transitions of wild-type apoMb (○) compared with the following mutants: S108L (●), F123W (□), F123G (■), and A130S (△). Conditions: 10 mM sodium acetate buffer, 0 °C, protein concentration 1.00 μM. Data are from Hughson et al. (1991, Fig. 2). Solid curves are drawn by inspection to aid viewing and are not the result of using the fitting procedure.

Table 1. *Sequences and ellipticities of A-, G-, and H-helix peptides*

Peptide	Sequence[a]	$[\theta]_{222}$[b]
A_{wt}	(Ac) SEGEWQLVLHVWAKV (NH$_2$)	Not helical
A_2	(Ac) SEGEAQLVLHVAAKVY (NH$_2$)	Not helical
G_{wt}	(Ac) PIKYLEFISEAIIHVLHSR (NH$_2$)	N/D
G_1	(Ac) PIKYLEEISEAIIKELHSK (NH$_2$)	−5,800
G_2	(Ac) PIKYLEELSEAIIKELHAK (NH$_2$)	−9,400
H_{wt}	(Ac) ADAQGAMNKALELFRKDIAAKYKE (NH$_2$)	−9,600
H_1	(Ac) ADAQEAMNKALELFRKDIAAKYKA (NH$_2$)	−15,500
ApoMb I		−15,000

[a] A_{wt} has a sequence identical to residues 3–18 of sperm whale myoglobin, G_{wt} is identical to residues 100–118, and H_{wt} is identical to residues 124–148. Sequence substitutions (bold) are designed to increase helix content, eliminate potential contributions of tryptophan to $[\theta]_{222}$ (A_2), and increase solubility (A_2, G_1, and G_2). Peptides were acetylated on their N-termini and amidated on their C-termini.

[b] $[\theta]_{222}$ was determined at 0 °C, pH 4.4 for A peptides, pH 4.2 for apoMb, and at 3 °C, pH 6.0 for G and H peptides. Values are in deg·cm^2·dmol^{-1}.

units of secondary structure in the native protein, although unstable in isolation, can be stabilized in early folding intermediates by interaction between their hydrophobic surfaces before fixed tertiary interactions are formed. This hypothesis can explain why native secondary structures are stabilized in molten globule intermediates even if no fixed tertiary interactions exist between different units of secondary structure: each unit of secondary structure typically has at least one hydrophobic face, and loose interaction between their hydrophobic surfaces will preferentially stabilize native secondary structures over nonnative structures that lack a continuous hydrophobic surface.

Concluding remarks

The molten globule model for the kinetics of protein folding applies when the molten globule is more stable than the unfolded form in refolding conditions and when α-helices in molten globule intermediates are stabilized by loose hydrophobic interactions between helices, which allows molten globules to be formed rapidly. Because molten globule intermediates contain stable α-helices at the locations of helices in the native protein, they provide a suitable matrix in which later folding steps can take place, resulting finally in a close-packed tertiary structure. The nature of these later steps will have to be the focus of much future work on understanding the kinetic process of folding.

The molten globule model for the kinetics of protein folding was stated convincingly for α-lactalbumin and has now been demonstrated for apoMb. Nevertheless, many important questions concerning the model remain unanswered. Does the molten globule model apply to proteins that do not show equilibrium molten globule intermediates? What are the topologies of helices in molten globule intermediates? Why are the B and E helices not

stabilized, as the A, G, and H helices are stabilized, in the equilibrium molten globule intermediate of apoMb? Are there a series of populated intermediates in the kinetic process of folding and, if so, where does the molten globule intermediate fit into the series? Is it possible for a β-sheet protein to form a molten globule intermediate with native-like secondary structure, resembling in this way the helical folding intermediates shown by cyt *c*, apoMb, and α-LA?

These and other unanswered questions show that work on the role of molten globule intermediates in the kinetic process of folding is only beginning.

Note added in proof

Stabilization of associated α-helices in a loose arrangement lacking close side-chain packing can be predicted from a calculation made by Richmond and Richards (1978: see their Fig. 4) that shows that loss of solvent accessibility of nonpolar side chains begins at helix-separation distances as great as 6 Å, a distance too large to permit side-chain packing.

Acknowledgments

The citation for the Stein and Moore Award refers to studies of kinetic intermediates on the pathway of protein folding. R.L.B. acknowledges the former students, fellows, and visitors who did this work, which began in 1971: Arlene Blum, David Brems, Kem Cook, Jean-Renaud Garel, Paul Hagerman, Fred Hughson, Magali Jullien, Peter Kim, Kunihiro Kuwajima, Alex Labhardt, Bob Matthews, Peter McPhie, Barry Nall, Drew Ridge, Andy Robertson, Franz Schmid, David Schultz, Tian Tsong, and Jayant Udgaonkar. He also thanks John Schellman for his discussion of the problems encountered in this work. Finally, R.L.B. gratefully acknowledges the support over many years of Virginia Robbins, who kept our laboratory running and also provided valuable technical help. D.B. acknowledges support from a Howard Hughes predoctoral fellowship. We thank Drs. F.W. Dahlquist, A.R. Fersht, and K. Kuwajima for permission to reproduce their published data. We thank Dr. Pat Jennings for a critical reading of this manuscript, for helpful discussions, and for communicating results prior to publication. This research was supported by NIH grant GM19988.

References

Baldwin, R.L. (1989). How does protein folding get started? *Trends Biochem. Sci. 14*, 291–294.

Baldwin, R.L. (1993). Pulsed H/D exchange studies of folding intermediates. *Curr. Opin. Struct. Biol. 3*, 84–91.

Barrick, D. & Baldwin, R.L. (1993). Three-state analysis of sperm whale apomyoglobin folding. *Biochemistry 32*, in press.

Baum, J., Dobson, C.M., Evans, P.A., & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: Studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry 28*, 7–13.

Brooks, C.L., III. (1992). Characterization of "native" apomyoglobin by molecular dynamics simulation. *J. Mol. Biol. 227*, 375–380.

Bycroft, M., Matouschek, A., Kellis, J.T., Jr., Serrano, L., & Fersht, A. (1990). Detection and characterization of a folding intermediate in barnase by NMR. *Nature 346*, 488–490.

Chothia, C., Levitt, M., & Richardson, D. (1981). Helix to helix packing in proteins. *J. Mol. Biol. 145*, 215–250.

Chyan, C.-L., Wormald, C., Dobson, C.M., Evans, P.A., & Baum, J. (1993). Structure and stability of the molten globule state of guinea pig α-lactalbumin: A hydrogen exchange study. *Biochemistry 32*, in press.

Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Yu., & Ptitsyn, O.B. (1981). α-Lactalbumin: Compact state with fluctuating tertiary structure? *FEBS Lett. 136*, 311–315.

Elöve, G.A., Chaffotte, A.F., Roder, H., & Goldberg, M.E. (1992). Early steps in cytochrome *c* folding probed by time-resolved circular dichroism and fluorescence spectroscopy. *Biochemistry 31*, 6876–6883.

Englander, S.W. & Kallenbach, N.R. (1984). Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys. 16*, 521–655.

Ewbank, J.J. & Creighton, T.E. (1991). The molten globule protein conformation probed by disulphide bonds. *Nature 350*, 518–520.

Flanagan, J.M., Kataoka, M., Shortle, D., & Engelman, D.M. (1992). Truncated staphylococcal nuclease is compact but disordered. *Proc. Natl. Acad. Sci. USA 89*, 748–752.

Goto, Y., Calciano, L.J., & Fink, A.L. (1990). Acid-induced folding of proteins. *Proc. Natl. Acad. Sci. USA 87*, 573–577.

Goto, Y. & Fink, A.L. (1990). Phase diagram for acidic conformational states of apomyoglobin. *J. Mol. Biol. 214*, 803–805.

Hughson, F.M., Barrick, D., & Baldwin, R.L. (1991). Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry 30*, 4113–4118.

Hughson, F.M., Wright, P.E., & Baldwin, R.L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science 249*, 1544–1548.

Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986a). Evidence for the identity between the equilibrium unfolding intermediate and a transient folding intermediate: A comparative study of the folding reactions of α-lactalbumin and lysozyme. *Biochemistry 25*, 6965–6972.

Ikeguchi, M., Kuwajima, K., & Sugai, S. (1986b). Ca^{2+}-induced alteration in unfolding behavior of α-lactalbumin. *J. Biochem (Jpn.) 99*, 1191–1201.

Jeng, M.-F., Englander, S.W., Elöve, G.A., Wand, A.J., & Roder, H. (1990). Structural description of acid-denatured cytochrome *c* by hydrogen exchange and 2D NMR. *Biochemistry 29*, 10433–10437.

Jennings, P.A. & Wright, P.E. (1992). NMR evidence that a molten globule intermediate is on the kinetic folding pathway for apomyoglobin. *Proc. Sixth Symp. Protein Soc., San Diego, California*, Abstr. M4.

Kuwajima, K. (1977). A folding model of α-lactalbumin produced from the three-state denaturation mechanism. *J. Mol. Biol. 114*, 241–258.

Kuwajima, K. (1989). The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. *Proteins Struct. Funct. Genet. 6*, 87–103.

Kuwajima, K., Nitta, K., Yoneyama, M., & Sugai, S. (1976). Three-state denaturation of α-lactalbumin by guanidine hydrochloride. *J. Mol. Biol. 106*, 359–373.

Lu, J. & Dahlquist, F.W. (1992). Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two dimensional NMR. *Biochemistry 31*, 4749–4756.

Ohgushi, M. & Wada, A. (1983). 'Molten globule state': A compact form of globular proteins with mobile side chains. *FEBS Lett. 164*, 21–24.

Ptitsyn, O.B. (1987). Protein folding: Hypotheses and experiments. *J. Protein Chem. 6*, 273–293.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E., & Razgulyaev, O.I. (1990). Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett. 262*, 20–24.

Radford, S.E., Dobson, C.M., & Evans, P.A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature 358*, 302–307.

Richmond, T.J. & Richards, F.M. (1978). Packing of α-helices: Geometrical constraints and contact areas. *J. Mol. Biol. 119*, 537–555.

Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, U.N., Gripas, A.F., & Gilmanshin, R.I. (1991). Study of the molten glob-

ule intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers 31*, 119–128.

Shortle, D. & Meeker, A.K. (1989). Residual structure in large fragments of staphylococcal nuclease: Effects of amino acid substitutions. *Biochemistry 28*, 936–944.

Takano, T. (1977). Structure of myoglobin refined at 2.0 Å resolution. I. Crystallographic refinement of metmyoglobin from sperm whale. *J. Mol. Biol. 110*, 537–568.

Tirado-Rives, J. & Jorgensen, W.L. (1993). Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry 32*, in press.

Waltho, J.P., Feher, V.A., Lerner, R.A., & Wright, P.E. (1989). Conformation of a T-cell stimulating peptide in aqueous solution. *FEBS Lett. 250*, 400–404.

Woodward, C., Simon, I., & Tüchsen, E. (1982). Hydrogen exchange and the dynamic structure of proteins. *Mol. Cell. Biochem. 48*, 135–160.

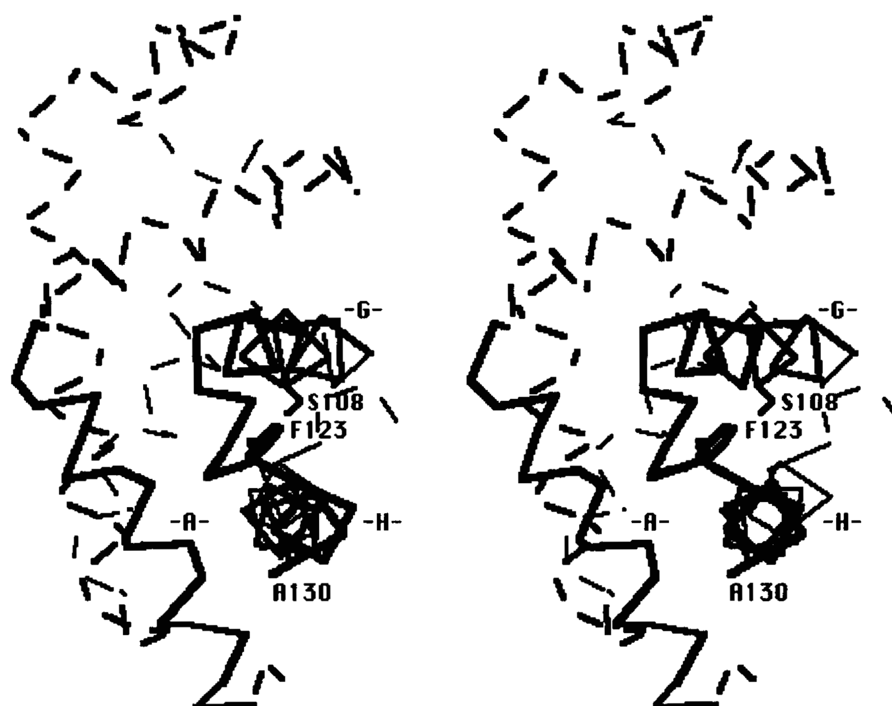

Figure added in proof. Helices A, G, and H of apomyoglobin (solid lines), which are present in an early folding intermediate. Side chains are shown and labeled for the three mutation sites at the helix interfaces. (Modified from kinemage file Barrick.kin, Diskette Appendix.)