Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 5446–5450, June 1995
Biochemistry

Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway

(protein folding/NMR/hydrogen exchange)

STEWART N. LOH, MICHAEL S. KAY, AND ROBERT L. BALDWIN†

Department of Biochemistry, Stanford University Medical Center, Stanford, CA 94305-5307

Contributed by Robert L. Baldwin, March 2, 1995

ABSTRACT **Apomyoglobin folding proceeds through a molten globule intermediate (low-salt form; I_1) that has been characterized by equilibrium (pH 4) and kinetic (pH 6) folding experiments. Of the eight α-helices in myoglobin, three (A, G, and H) are structured in I_1, while the rest appear to be unfolded. Here we report on the structure and stability of a second intermediate, the trichloroacetate form of the molten globule intermediate (I_2), which is induced either from the acid-unfolded protein or from I_1 by ≥5 mM sodium trichloroacetate. Circular dichroism measurements monitoring urea- and acid-induced unfolding indicate that I_2 is more highly structured and more stable than I_1. Although I_2 exhibits properties closer to those of the native protein, one-dimensional NMR spectra show that it maintains the lack of fixed side-chain structure that is the hallmark of a molten globule. Amide proton exchange and ^1H–^{15}N two-dimensional NMR experiments are used to identify the source of the extra helicity observed in I_2. The results reveal that the existing A, G, and H helices present in I_1 have become more stable in I_2 and that a fourth helix—the B helix—has been incorporated into the molten globule. Available evidence is consistent with I_2 being an on-pathway intermediate. The data support the view that apomyoglobin folds in a sequential fashion through a single pathway populated by intermediates of increasing structure and stability.**

As information regarding early events in protein folding becomes available, attention is being directed towards the enigmatic state known as the molten globule. Molten globules appear as common intermediates in the kinetic folding process of several proteins (1–4), and it has been suggested that they may be general folding intermediates (5, 6). Their structures and the forces that underlie their stability are essential—but largely missing—pieces of the folding puzzle. Apomyoglobin [apoMb; sperm whale myoglobin (Mb) that lacks the heme group] is well suited to address these key issues. Its folding is known to proceed through a molten globule intermediate (low-salt form; I_1) that is observed in equilibrium (pH 4) (7, 8) and kinetic (pH 6) (9) folding experiments. The I_1 state exhibits properties of a typical molten globule: high secondary structure content as measured by CD spectroscopy, yet little fixed side-chain structure as determined by ^1H NMR experiments (i.e., minimal dispersion of side-chain resonances; see *Results*). Based on the pattern of exchange-protected amide protons, folding of I_1 is largely associated with a subdomain formed by the A, G, and H helices; the remaining helices are disordered (7). Interest in this equilibrium molten globule is especially high because it satisfies some requirements for being an on-pathway folding intermediate. These requirements are: (*i*) that it be stable relative to the unfolded protein (U) under folding conditions (8), (*ii*) that it be formed quickly from U (9), and (*iii*) that its conversion to the native protein (N) and U be completely reversible (7).

Although it remains difficult to characterize molten globules, it is clear that they are fundamentally unlike their corresponding native states, with many of the close-packed interactions that stabilize N being lessened or completely absent (10). Important questions thus become, at what point in folding are native-like side-chain interactions gained, and what causes this event? Goto, Fink, and coworkers have shown that the addition of certain anions to acid-unfolded proteins can trigger partial or total refolding (11). When a partly folded structure is formed, it exhibits properties that define it to be a molten globule. apoMb is an example in which a molten globule (I_1) may be formed either by this route (anion binding to the acid-denatured form at pH 2) or by pH-induced unfolding at an intermediate pH (pH 4), at very low salt concentrations (12). Curiously, addition of CCl_3COONa results in a significant increase in helicity (13).

Here we employ CD spectroscopy and amide proton exchange measurements to examine the folding of apoMb in the presence of 20 mM CCl_3COONa. With CCl_3COONa present, the equilibrium folding intermediate is not I_1 but rather is a new species, I_2, that is more highly structured and more stable relative to U. Our goal is to characterize the structure of this new partly folded species and to determine its place in a possible hierarchy of folding intermediates. The identity of I_2 is established by NMR-monitored hydrogen exchange. The locations of protected protons are similar in I_2 and in an intermediate formed immediately after I_1 in the kinetic pathway of apoMb folding (pH 6 and in the absence of CCl_3COONa), in which the B helix is stabilized (9). These results support the view that folding proceeds sequentially through a series of intermediates, each with progressively higher stability and structure content.

MATERIALS AND METHODS

Protein Preparation. To obtain high levels of isotopically labeled protein, the Mb gene from the plasmid pMb413b [originally a gift from S. G. Sligar (University of Illinois) (14) and modified as described (10)] was incorporated into the phage T7 vector pT7-7 (United States Biochemical) in the following way. An *Nde* I site, corresponding to Met-1 of the Mb sequence, was introduced into pMb413b by using the polymerase chain reaction. The gene was then excised by digestion with *Sal* I and *Nde* I and ligated to the *Sal* I–*Nde* I fragment of pT7-7 to generate the new plasmid pMbT7. The entire Mb sequence was verified by standard dideoxy methods.

^{15}N-labeled protein was overexpressed by using the *Escherichia coli* strain BL21 (DE3) (Novagen), grown in minimal M9

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Mb and apoMb, sperm whale myoglobin with and without the heme group, respectively; I_1 and I_2, low-salt form and trichloroacetate form, respectively, of the molten globule intermediate of apoMb; U, unfolded protein; N, native protein.

†To whom reprint requests should be addressed.

medium (15) containing 1 g of $(^{15}NH_4)_2SO_4$ (99 atom % enriched) (Isotec, Miamisburg, OH) per liter. Cultures were grown at 37°C in a New Brunswick fermentor and induced at A_{600} = 1.0 with 25 mg of dioxane-free isopropyl β-D-thiogalactopyranoside (Sigma) per liter. After an additional 6 hr at 37°C, the cells were harvested by centrifugation and lysed by the lysozyme method (16). Most of the protein produced was in the apo form and insoluble; these inclusion bodies were broken up by the addition of 10 mM KCN, a slight excess of hemin dicyanide and subsequent sonication. In this way the apoMb was converted to the highly soluble cyano-MetMb, which contains the heme group but with iron in the oxidized 3+ state. The lysate was desalted by dialysis against 10 mM Tris acetate (pH 5.5) and then passed through sequential DE 52 and CM 52 columns. Elution from the latter column was by means of a 0 to 0.5 M NaCl gradient (pH 5.5). After dialysis against H_2O, the protein was judged to be >95% pure by SDS gel electrophoresis; yields were ≈50 mg/liter. Removal of the heme group was by the acid-acetone method (17).

Hydrogen Exchange. All steps were carried out in a 4°C cold room. Exchange was initiated by a 10-fold dilution of 0.8 mM apoMb (pH 6.0 in H_2O) into 25 ml of 10 mM [2-$^2H_3$]acetate (pH* 4.2, ^2H_2O; pH* is the glass electrode reading unadjusted for the deuterium isotope effect) either in the presence or absence of 20 mM CCl_3COONa. After a variable time, the exchange reaction was quenched by addition of 1/10th volume of 1 mM hemin dicyanide (pH* 9). Hemin dicyanide was used because it remains monomeric at high concentrations (18). A previously determined amount of 0.5 M KCN was simultaneously added to raise the final pH* to 5.6; under these conditions the protein rapidly refolds to generate cyano-MetMb. After concentration to approximately 2 ml (Amicon YM10 membrane), the solution was passed through a Pharmacia PD10 column (equilibrated with 10 mM [2-$^2H_3$]acetate, ^2H_2O, pH* 5.6) to remove excess solutes and residual H_2O. Cyano-MetMb was converted to MbCO by reduction with 0.01 volume of 0.5 M $Na_2S_2O_4$, which was added to the CO-saturated protein solution with stirring and under constant pressure of CO. Residual $Na_2S_2O_4$ and CN^- were effectively removed by repeated dilution with CO-saturated 10 mM [2-$^2H_3$]acetate (pH* 5.6, ^2H_2O) and ultrafiltration under helium gas. Samples were concentrated to 0.5 ml (1–2 mM Mb) and incubated at 35°C for 12 hr immediately prior to NMR data collection.

NMR Spectroscopy. NMR spectra were recorded with a 500 MHz General Electric GN-Omega spectrometer. Data were collected at 35°C unless otherwise noted. Proton occupancies at each backbone amide position were quantitated by means of ^1H–^{15}N heteronuclear multiple-quantum correlation (HMQC) spectra collected in hypercomplex mode (19, 20). Two-dimensional data sets consist of 256 t_1 increments of 1024 complex t_2 data points, covering spectral widths of 1667 Hz and 6006 Hz in the ^{15}N and ^1H dimensions, respectively. ^{15}N decoupling was achieved by a WALTZ-16 sequence (21) applied during acquisition. Prior to Fourier transformation, the data were zero-filled to 2048 × 1024 complex points (t_2 × t_1), and shifted sine-bell squared functions were applied to each dimension. Resonance assignments were obtained from ref. 22.

Cross peak volumes were calculated and optimized by using the FELIX software package (release 2.3, Biosym Technologies, San Diego). Volumes within a given spectrum were normalized to the area of the Val-68 $C^\gamma H_3$ peak at −2.3 ppm, averaged over one-dimensional ^1H spectra taken before and after each heteronuclear multiple-quantum correlation experiment. Amide proton exchange rates (k_{obs}) were obtained from single exponential fits of the decaying NH cross peak volumes as a function of exchange time. Intrinsic exchange rates for unprotected, solvent-exposed protons (k_{exp}) were calculated from model peptide data (23). The results are reported as the protection factor $P = k_{exp}/k_{obs}$.

CD Measurements. Data were collected on an Aviv 60DS spectropolarimeter employing a thermostat-controlled sample holder set to 4°C. For samples prepared without CCl_3COONa, a protein concentration of 2 μM and a 10-mm path-length cuvette were used. Samples containing 20 mM CCl_3COONa required a 1-mm path-length cuvette and an apoMb concentration of 15 μM. The buffer was 3 mM citrate.

RESULTS

Structure and Stability of Intermediates Monitored by CD. The presence of intermediates is revealed by the three-state character of acid-induced unfolding (Fig. 1). I_1, which is maximally populated at pH 4.2, exhibits a $[\theta]_{222}$ value of −14,500, corresponding to a helix content 70% that of N. Addition of 5 mM CCl_3COONa increases the helicity of the intermediate and stabilizes it relative to U, as evidenced by the large shift in the pH midpoint of the $I_2 \rightleftharpoons U$ transition to lower pH (Fig. 1). Doubling the CCl_3COONa concentration eliminates acid unfolding of I_2 altogether (data not shown) and further increases its helix content, although the greatest change in ellipticity occurs between 0 and 5 mM CCl_3COONa (Fig. 1 *Inset*). Aggregation of apoMb is a well-known problem (24, 25) that can potentially influence the CD signal as well as the amide proton exchange behavior. To test the effect of protein concentration on the CD signal, $[\theta]_{222}$ values were recorded as a function of apoMb concentration and were found to remain constant up to 80 μM apoMb at pH 4 and at pH 6, in both the presence and absence of 20 mM CCl_3COONa (data not shown). The acid-induced $I_2 \rightleftharpoons N$ and $I_2 \rightleftharpoons U$ transitions are reversible by the test that the same transition curves are obtained by unfolding (pH jumps starting from pH 6) and by refolding (pH jumps starting from pH 2).

The large difference in stabilities of I_1 and I_2 is illustrated by the urea-induced unfolding curve shown in Fig. 2. The I_2 unfolding data are well fitted to a two-state model (26), yielding the following parameters: $\Delta G° = -5.9$ kcal·mol^{-1}, C_m = 3.9 M, m = 1520 cal·mol^{-1}·M^{-1}, where $\Delta G°$ is the standard change in Gibbs energy at 0 M urea, C_m is the urea molarity at the midpoint of the unfolding transition, and m is the slope of $\Delta G°$ versus molar urea. I_2 is remarkably stable, considering that a previous study reports $\Delta G° = -6.6$ kcal·mol^{-1} for $N \rightleftharpoons U$ and $\Delta G° = -2.1$ kcal·mol^{-1} for $I_1 \rightleftharpoons U$ (8). The same study,

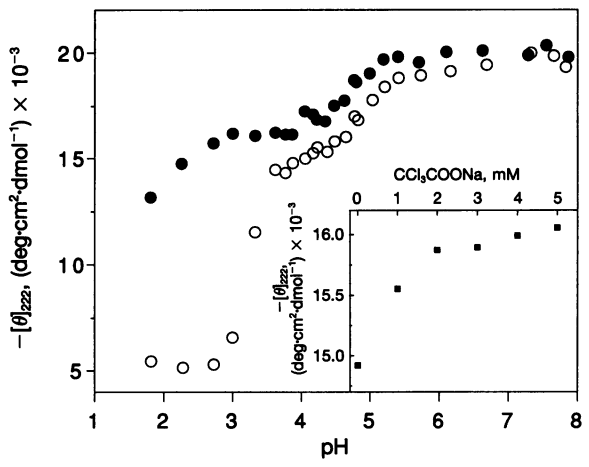

FIG. 1. Acid-induced unfolding of apoMb in 20 mM CCl_3COONa (●) and in the absence of CCl_3COONa (○), monitored by the molecular ellipticity at 222 nm ($[\theta]_{222}$). The three-state character of unfolding illustrates the presence of equilibrium intermediates in both cases. All transitions are fully reversible. (*Inset*) CCl_3COONa titration of the equilibrium molten globule intermediate at pH 4.2.

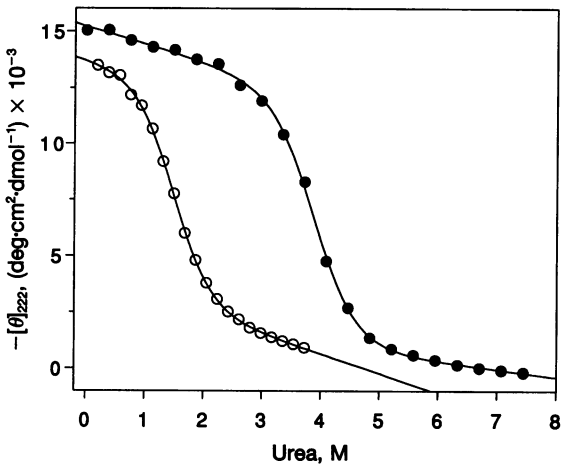

FIG. 2. Urea-induced unfolding of I_1 (no CCl_3COONa present) (○) and I_2 (20 mM CCl_3COONa present) (●) at pH 4.2 as monitored by CD spectroscopy. Lines represent the best fit of the data to a two-state unfolding model.

which used simultaneous three-state fits of combined urea- and acid-unfolding curves, finds the m values for the $N \rightleftharpoons I_1$ and $I_1 \rightleftharpoons U$ transitions to be 1030 cal·mol^{-1}·M^{-1} and 1010 cal·mol^{-1}·M^{-1}, respectively. To determine whether protein aggregation contributes to the apparent stability of I_2, the experiments depicted in Fig. 2 were repeated with apoMb concentrations ranging from 10 μM to 100 μM; the resulting curves were superimposable within experimental error (not shown). Unfolding of I_1 can be fitted to a two-state model (Fig. 2). The $I_1 \rightleftharpoons U$ parameters are $\Delta G^\circ = -2.5$ kcal·mol^{-1}, m = 1720 cal·mol^{-1}·M^{-1}, C_m = 1.5 M, although the poorly defined baselines prevent these values from being determined accurately (urea concentrations > 4 M are inaccessible without introducing HCl to adjust pH).

To address the possibility that I_2 might be a native-like intermediate or the native species itself, one-dimensional ^1H spectra were recorded under unfolded, native, and intermediate conditions (Fig. 3). Close-packed tertiary structure, which gives rise to local magnetic microenvironments, is manifest in NMR spectra by up-field shifting of methyl peaks, dispersion in the aromatic region, and spreading out of the histidine $^1H^{\varepsilon 1}$ resonances. All these characteristics are present

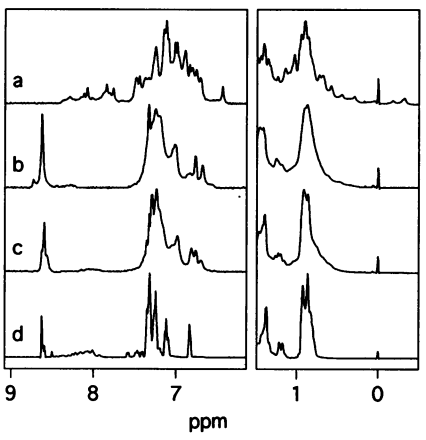

FIG. 3. The different states of apoMb characterized by ^1H NMR spectroscopy. Traces: a, N at pH 6.1; b, I_2 at pH 4.2 in 20 mM CCl_3COONa; c, I_1 at pH 4.2; and d, U at pH 2.3. Shown are the histidine resonances of the $^1H^{\varepsilon 1}$ region (7.7–8.7 ppm), aromatic region (6.5–7.5 ppm), and upfield-shifted methyl region (−0.5–1.0 ppm). Sample conditions are 110 μM apoMb and 10 mM [2-$^2H_3$]acetate in ^2H_2O at 20°C.

in N and largely absent in U, I_1, and I_2, although the slight dispersion in the aromatic region of I_1 and I_2 suggests that small amounts of side-chain structure exist in both species. Subtle changes in ^1H spectra further hint at differences in conformation between I_1 and I_2.

Hydrogen Exchange of Intermediates. The two-dimensional ^1H–^{15}N spectrum taken at zero time of exchange (not shown) reveals that ≈50 NH protons exchange slowly enough to serve as structural probes. These probes are distributed among the A, B, C, E, G, and H helices; no information is available regarding the participation of helices D and F in the molten globule. Fig. 4 depicts a representative plot of the loss of proton occupancy as a function of time. As reflected by the difference in decay times, hydrogen exchange provides a detailed and sensitive method for detecting both changes in structure and stability. Exchange data, expressed in the form of protection factors, are summarized in Table 1. Fig. 5 depicts residues with protection factors > 5 in the I_2 state, as shown by their locations in the native Mb x-ray crystal structure.

DISCUSSION

Nature of the Folding Pathway. Recent theoretical (refs. 28 and 29 and references therein) and experimental (30) studies have challenged the conventional view that proteins fold by unique pathways with discrete and sequentially populated intermediates. Statistical mechanical theories and Monte Carlo simulations suggest that folding may occur instead through a continuum of states and pathways and that experimentally observed intermediates are nonobligatory (refs. 28 and 29; see also ref. 31). This view holds that folding intermediates are in fact kinetically trapped and misfolded species. Cytochrome *c* provides experimental evidence for such "fast track" folding. When a nonnative interaction between a histidine residue and the heme group is prevented from being formed, folding occurs rapidly and without observable kinetic intermediates (30). We show here that two different equilibrium intermediates are found in the folding of apoMb; they possess various amounts of native-like structure and exhibit correspondingly different stabilities. If it can be shown that they satisfy the requirements for being on-pathway, this will constitute strong support for a sequential folding pathway and provide a system for delineating its mechanistic details.

Addition of small amounts of CCl_3COONa to I_1 (or U) results in a species with higher secondary structure content and enhanced stability. CCl_3COONa has no observable effect on

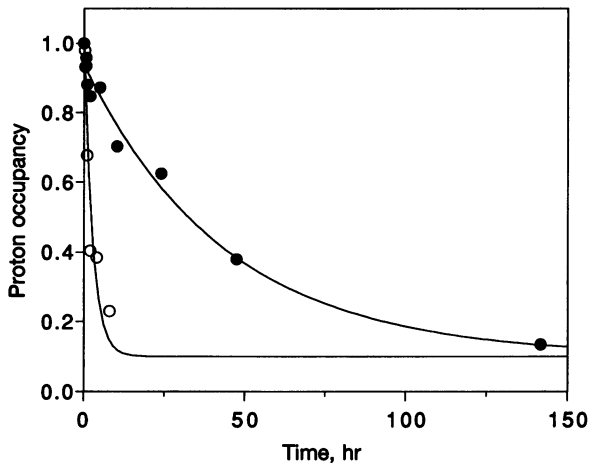

FIG. 4. Representative plot of proton occupancy as a function of exchange time for the B-helix residue Ile-30 in the I_1 state (no CCl_3COONa present) (○) and in the I_2 state (20 mM CCl_3COONa present) (●). Lines are the best fit of the data to single exponential decays.

Table 1. Protection factors of I_1 and I_2

Residue	Position	$P(I_1)$	$P(I_2)$	$P(I_2)/P(I_1)$
Leu-9	A7	0.44	7.8	18
Leu-11	A9	12	660	55
Trp-14	A12	7.2	410	57
Val-17	A15	4.6	260	56
Glu-18	A16	2.0	32	16
Asp-27	B8	0.60	1.6	2.7
Ile-28	B9	0.62	0.62	1.0
Leu-29	B10	1.4	12	8.5
Ile-30	B11	1.8	43	24
Arg-31	B12	1.2	13	11
Leu-32	B13	1.2	150	120
Phe-33	B14	4.0	34	8.4
Lys-34	B15	1.1	5.4	4.9
Thr-39	C4	0.74	1.0	1.5
Leu-40	C5	0.40	0.93	2.3
Val-66	E9	0.30	0.45	1.5
Val-68	E11	1.0	0.48	0.48
Thr-70	E13	0.45	0.74	1.6
Ala-71	E14	1.9	1.3	0.66
Leu-72	E15	0.30	0.58	1.9
Gly-73	E16	0.90	1.1	1.2
Ala-74	E17	1.0	1.4	1.4
Ile-75	E18	0.30	0.73	2.4
Lys-76	E19	0.22	0.50	2.2
Lys-77	E20	0.38	0.54	1.4
Tyr-103	G4	0.61	6.3	10
Leu-104	G5	1.3	43	33
Ile-107	G8	3.1	410	130
Ala-110	G11	4.1	220	53
Ile-112	G13	5.5	410	74
Val-114	G15	4.5	220	50
Leu-115	G16	1.0	49	49
Ala-134	H11	7.6	44	5.8
Leu-135	H12	8.9	120	13
Glu-136	H13	15	110	7.5
Leu-137	H14	6.4	89	14
Arg-139	H16	12	620	52
Lys-140	H17	2.5	45	18
Ile-142	H19	5.0	110	22
Ala-143	H20	4.5	180	41

the secondary structure of the native state (Fig. 1) or on its stability as determined from urea-induced unfolding curves (not shown). The striking difference in stability between I_1 and I_2 is revealed in Fig. 2. Unfolding of I_2 is apparently cooperative and can be well approximated by a two-state transition. Standard analysis indicates that I_2 is only slightly less stable than N. In contrast, I_1 is marked by a significantly lower midpoint of urea denaturation. There are conflicting views of whether the $I_1 \rightleftharpoons U$ transition is two-state (13, 32). A lower limit on its stability, imposed by fitting to a two-state model, yields a standard free energy that is less than half that of I_2 (Fig. 2 and ref. 8). These data clearly demonstrate that I_2 exhibits properties closer to those of the native state than does I_1. Nevertheless, I_2 maintains the lack of fixed side-chain structure that is diagnostic of a molten globule (Fig. 3).

Structure and Properties of I_2. What is the source of the extra helicity observed in I_2, and which elements of structure are stabilized relative to I_1? We first consider the hydrogen exchange behavior of I_1. All significantly protected residues (i.e., those with protection factors > 5) map to the A, G, and H helices. The locations of protected amide protons are in qualitative agreement with those previously reported (7) and suggest that secondary structure in I_1 is associated exclusively with portions of the A–G–H subdomain. Many of the I_1 protection factors are significantly lower, however, in the present study. This discrepancy can be attributed in large part to the 75% lower intrinsic exchange rates of valine, leucine, and isoleucine used in computing the protection factors in Table 1 (23, 33). The remaining inconsistency may be caused by the different quenching conditions used in the two studies. Overall, I_1 appears to be less stable than previously thought, with a maximum A–G–H protection factor of 15. The B, C, and E helices are completely unstructured.

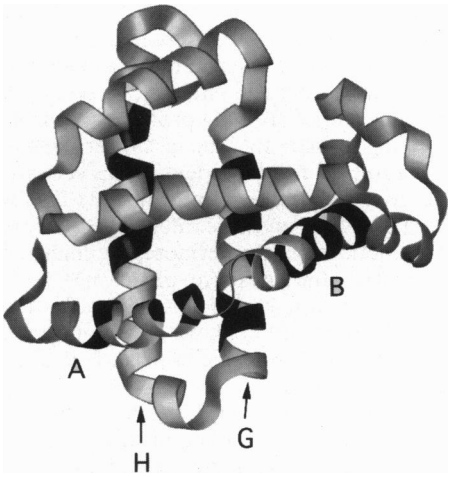

FIG. 5. The locations of significantly protected residues in I_2 ($P \geq 5$) superimposed on the x-ray crystal structure of MbCO (27). The A, B, G, and H helices are indicated. The INSIGHT II software package (release 2.3, Biosym Technologies) was used to represent the structure.

NMR-monitored hydrogen exchange data permit one to analyze changes in stability and structure at the level of individual amide groups. The CCl_3COONa-induced changes in I_2, detected by CD experiments monitoring acid- and urea-induced unfolding, can now be attributed to two effects. First, the overall structure of I_1 is significantly stabilized in I_2. Protection factors of residues in the A–G–H subdomain are as high as 660, up by an average factor of 40. Second, the previously unstructured B helix acquires substantial protection in I_2, indicating that it has become folded. As in I_1, the C and E helices are essentially unprotected. Thus, the apoMb I_2 state consists of stabilized A, B, G, and H helices and unstructured C and E helices.

We now suggest a more detailed structural model for I_2 based on the exchange data. All available probes in the A–G–H subdomain are significantly protected, and each of these helices contains residues with protection factors in excess of 400. For the most part, residues in the B helix are less protected than those in the A, G, and H helices, and several are unprotected. Furthermore, protection factors in the B helix only increase by an average factor of 20 as a result of CCl_3COONa addition. One structure that is consistent with these observations consists of a stable A–G–H subdomain, more closely packed than its marginally stable counterpart in I_1, with the B helix loosely associated with it—perhaps beginning to assemble onto this nascent structure. If this is correct, then the arrangement of helices in I_2 is likely to be different than in native apoMb because several of the protected B-helix residues in I_2 are spatially removed from the A–G–H subdomain in holoMb (Fig. 5). If hydrogen exchange in I_2 is limited by global unfolding and occurs by the EX2 mechanism, then protection factors of $\approx 10^4$ should be observed, given the value of $\Delta G°$ found from Fig. 2. The largest protection factor is in fact lower by a factor of 20. Therefore, exchange by mechanisms other than global unfolding, probably like those found in native proteins, must contribute to exchange in I_2.

The protected amide protons of the I_1 equilibrium molten globule share an almost exact correspondence with those of an early kinetic intermediate that is formed within the 5-ms dead time of rapid mixing experiments (9). A second species, in which all B helix probes are fully protected, appears after ≈1 s; the remaining helices fold on a slower time scale. The locations of protected residues identify this second species as being similar to the equilibrium intermediate I_2. Since folding and unfolding of I_2 to N and U is freely reversible, I_2 is not the product of a kinetic trap. Furthermore, the high stability of I_2 (at pH 4.2 and in the presence of 20 mM CCl_3COONa) indicates that it is also stable relative to U under native conditions (i.e., pH 6), although we have not directly demonstrated this point. By these three criteria, I_2 is likely to be an on-pathway intermediate.

What stabilizes I_2 relative to I_1? In general, the effectiveness of an anion in triggering refolding of acid-unfolded proteins follows the electroselectivity series—i.e., $SO_4^{2-} > CCl_3COO^- > ClO_4^- > NO_3^- > CF_3COO^- > Cl^-$ (12). Together with the fact that only low salt concentrations are needed to induce refolding (<5 mM, Fig. 1 and ref. 13), this finding strongly suggests that the mechanism by which CCl_3COONa stabilizes I_2 is charge neutralization by site binding to positive groups. An earlier study, which established an inverse relationship between the stability of the cytochrome *c* molten globule and the number of groups contributing to positive charge repulsion, provides a basis for this view (34). These data suggest that apoMb folding at pH 4 is arrested at the I_1 stage because unfavorable charge repulsion between positive groups prevents I_2 (and hence N) from being formed. It seems paradoxical that CCl_3COO^- (or some other appropriate anion) is needed to form I_2 although CCl_3COO^- has little effect on the stability of N. This paradox suggests that the factors driving and opposing folding are actually balanced quite differently in I_2 and N, even though both folded forms have similar stabilities.

Concluding Remarks. Our results add to the evidence that the folding pathway of apoMb can be described by the following minimal mechanism:

$$U \rightarrow I_1 \rightarrow I_2 \rightarrow N.$$

Structure in I_1 is limited to a marginally stable A–G–H subdomain, while I_2 consists of a more stable A–G–H subdomain associated with a loosely stable B helix. I_2 attains native-like stability without extensive close-packed side-chain interactions, which are formed relatively late in the folding pathway. apoMb provides support for the simple sequential model for protein folding. Critical experiments are now needed to determine if I_1 and I_2 are on-pathway folding intermediates.

We thank Dr. Jeung-Hoi Ha for the gift of the pT7-7 vector. S.N.L. is a fellow of the Damon Runyon–Walter Winchell Cancer Fund (DRG 1229). M.S.K. is a trainee of the National Institute of General Medical Sciences Medical Scientist Training Program (2T32GM07365). This work was supported by National Institutes of Health Grant GM 19988.

1. Kuwajima, K. (1977) *J. Mol. Biol.* **114,** 241–258.
2. Chen, B.-L., Baase, W. A., Nicholson, H. & Schellman, J. A. (1992) *Biochemistry* **31,** 1464–1476.
3. Ogasahara, K. & Yutani, K. (1994) *J. Mol. Biol.* **236,** 1227–1240.
4. Sanz, J. M., Johnson, C. M. & Fersht, A. R. (1994) *Biochemistry* **33,** 11189–11199.
5. Ptitsyn, O. B. (1987) *J. Protein Chem.* **6,** 273–293.
6. Kuwajima, K. (1989) *Proteins Struct. Funct. Genet.* **6,** 87–103.
7. Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990) *Science* **249,** 1544–1548.
8. Barrick, D. & Baldwin, R. L. (1993) *Biochemistry* **32,** 3790–3796.
9. Jennings, P. A. & Wright, P. E. (1993) *Science* **262,** 892–896.
10. Hughson, F. M., Barrick, D. & Baldwin, R. L. (1991) *Biochemistry* **30,** 4113–4118.
11. Goto, Y., Calciano, L. J. & Fink, A. L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 573–577.
12. Goto, Y., Takahashi, N. & Fink, A. L. (1990) *Biochemistry* **29,** 3480–3488.
13. Nishii, I., Kataoka, M., Tokunaga, F. & Goto, Y. (1994) *Biochemistry* **33,** 4903–4909.
14. Springer, B. A. & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8961–8965.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
16. Cull, M. & McHenry, C. S. (1990) *Methods Enzymol.* **182,** 147–154.
17. Fanelli, A. R., Antonini, E. & Caputo, A. (1958) *Biochim. Biophys. Acta.* **30,** 608–615.
18. Yee, S. & Peyton, D. H. (1991) *FEBS Lett.* **290,** 119–122.
19. Müller, L. (1979) *J. Am. Chem. Soc.* **101,** 4481–4484.
20. States, D. J., Haberkorn, R. A. & Rubin, D. J. (1982) *J. Magn. Reson.* **48,** 286–292.
21. Shaka, A. J., Keeler, J., Frenkiel, T. & Freeman, R. (1983) *J. Magn. Reson.* **53,** 335–338.
22. Thériault, Y., Pochapsky, T. C., Dalvit, C., Chiu, M. L., Sligar, S. G. & Wright, P. E. (1994) *J. Biomol. NMR* **4,** 491–504.
23. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993) *Proteins Struct. Funct. Genet.* **17,** 75–86.
24. Rumen, N. M. & Apella, E. (1962) *Arch. Biochem. Biophys.* **97,** 128–133.
25. De Young, L. R., Dill, K. A. & Fink, A. L. (1993) *Biochemistry* **32,** 3877–3886.
26. Santoro, M. M. & Bolen, D. W. (1988) *Biochemistry* **27,** 8063–8068.
27. Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. (1986) *J. Mol. Biol.* **192,** 133–154.
28. Šali, A., Shakhnovich, E. I. & Karplus, M. (1994) *Nature (London)* **369,** 248–251.
29. Onuchic, J. N., Wolynes, P. G., Luthy-Schulten, Z. & Socci, N. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3626–3630.
30. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994) *Nat. Struct. Biol.* **1,** 149–156.
31. Baldwin, R. L. (1995) *J. Biol. NMR* **5,** 103–109.
32. Griko, Y. V. & Privalov, P. L. (1994) *J. Mol. Biol.* **235,** 1318–1325.
33. Robertson, A. D. & Baldwin, R. L. (1991) *Biochemistry* **30,** 9907–9914.
34. Goto, Y. & Nishikiori, S. (1991) *J. Mol. Biol.* **222,** 679–686.