135. M. Baba *et al.*, *Antimicrob. Agents Chemother.* **31**, 1613 (1987); V. A. Johnson *et al.*, *J. Infect. Dis.* **159**, 837 (1989); S. Hayashi *et al.*, *Antimicrob. Agents Chemother.* **34**, 82 (1990); V. A. Johnson, M. A. Barlow, D. P. Merrill, T. C. Chou, M. S. Hirsch. *J. Infect. Dis.* **161**, 1059 (1990); M. W. Vogt *et al.*, *Science* **235**, 1376 (1987).
136. Unless shown otherwise, the in vitro antiviral activity of these compounds was determined by H. Mitsuya and S. Broder.
137. We thank T. Shimada and M. S. Reitz for critical reading of the manuscript.

Research Article

Structural Characterization of a Partly Folded Apomyoglobin Intermediate

FREDERICK M. HUGHSON, PETER E. WRIGHT, ROBERT L. BALDWIN*

To understand why proteins adopt particular three-dimensional structures, it is important to elucidate the hierarchy of interactions that stabilize the native state. Proteins in partly folded states can be used to dissect protein organizational hierarchies. A partly folded apomyoglobin intermediate has now been characterized structurally by trapping slowly exchanging peptide NH protons and analyzing them by two-dimensional ^{1}H-NMR (nuclear magnetic resonance). Protons in the A, G, and H helix regions are protected from exchange, while protons in the B and E helix regions exchange freely. On the basis of these results and the three-dimensional structure of native myoglobin, a structural model is presented for the partly folded intermediate in which a compact subdomain retains structure while the remainder of the protein is essentially unfolded.

THE FOLDING OF SMALL SINGLE-DOMAIN PROTEINS IS, IN most cases, highly cooperative. Stable, partly folded intermediates that might lend insight into the folding process are seldom observed. Intermediates are sometimes present transiently during folding, and techniques have been developed recently that probe their structures (*1, 2*). Such structural studies are technically demanding, and the existence of multiple unfolded forms adds to the difficulty of interpretation.

The failure to detect stable folding intermediates can be blamed on the stability of the native protein; conditions severe enough to unfold the native state usually destabilize partly folded states as well. Destabilizing the native state selectively, by removing a stabilizing ion or a prosthetic group or by site-directed mutagenesis, can promote the appearance of partly folded states. The best characterized example of a protein exhibiting such an "equilibrium" intermediate is α-lactalbumin, which in the presence of a bound Ca^{2+} unfolds cooperatively in a two-state reaction. When, however, the native state is destabilized by removal of the Ca^{2+}, the protein begins to unfold at a lower denaturant concentration, revealing a partly folded state (*3*).

Myoglobin (Mb) has proved to be a tractable protein for structural studies of a stable intermediate. Removing the heme group destabilizes native Mb and leads to the appearance of a stable intermediate in low pH–induced unfolding (*4*). At neutral pH, apo-Mb is compact with an α-helix content of ~55 percent as measured by circular dichroism (CD). Although the detailed three-dimensional structure of apo-Mb is not known, we show here that it resembles holo-Mb by criteria based on amide proton exchange. Others have shown (*4*) that apo-Mb contains an ordered hydrophobic core, as judged by calorimetric criteria. As the pH is lowered below neutral, "native" apo-Mb unfolds in two distinct stages (Fig. 1): first to a compact intermediate with a helix content of ~35 percent, termed I, and then to a more fully denatured state with little residual helix. The intrinsic viscosity of apo-Mb at pH 2 is large (*4*), as expected for a random-coil polypeptide chain.

The simple architecture of Mb, consisting of eight α helices that assemble to form a hydrophobic pocket for the heme group, has guided thinking about plausible pathways for protein folding. For example, one model postulates that helices form early in folding and are stabilized by subsequent helix pairing reactions (*5*). The discovery of a stable intermediate in low pH–induced unfolding offers the opportunity to study a partly folded state of apo-Mb in detail. Ultimately, however, kinetic experiments seem necessary to establish whether or not this intermediate lies on the folding pathway.

The apo-Mb I form tends to aggregate in solution (*6*). This property, although unsurprising because of the exposure of hydrophobic patches that might accompany partial unfolding, makes direct structure determination by x-ray crystallography or NMR difficult. Therefore, we used an indirect method to gather detailed structural information about the compact intermediate. This method allows amide proton exchange rates to be measured at about 40 sites distributed throughout the protein. Because the exchange of amide protons for solvent protons can be strongly retarded by hydrogen-bonded secondary structure, measurement of exchange rates serves to map out regions of the protein that are folded. In the I form, the regions of apo-Mb corresponding to the A, G, and H helices contain many protected protons, whereas protons in the B and E helix regions exchange at rates near those expected for an

F. M. Hughson and R. L. Baldwin are in the Department of Biochemistry, Beckman Center, Stanford University School of Medicine, Stanford, CA 94305. P. E. Wright is in the Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037.

*To whom correspondence should be addressed.

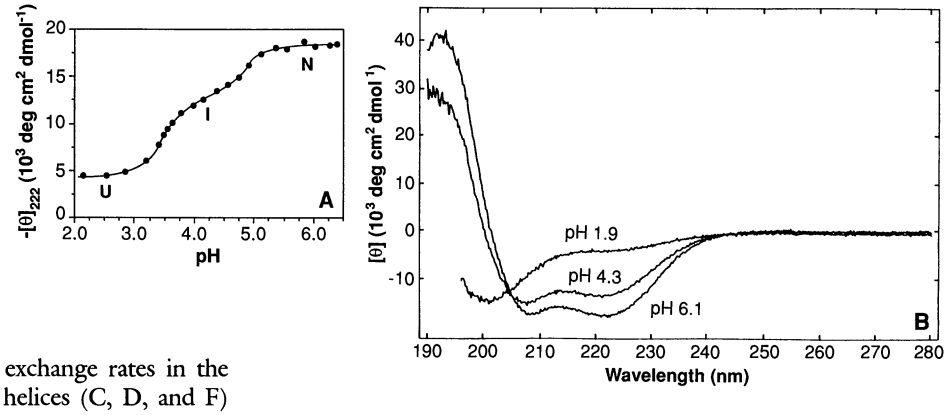

Fig. 1. (**A**) Apo-Mb unfolds in two discrete stages as the pH is lowered, as monitored by the mean residue ellipticity [θ] at 222 nm (*30*). Plateaus at high and low pH represent the native (N) and unfolded (U) states, respectively, while the sloping plateau centered around pH 4 represents the stable intermediate form. Cooperative transitions are seen among the three states. Both transitions are completely reversible. (**B**) The CD spectra of apo-Mb at pH 6.1, 4.3, and 1.9 (*30*). The spectral features observed at pH 6.1 and 4.3 are characteristic of partial α-helix formation (*31*).

unfolded polypeptide chain. Amide proton exchange rates in the regions corresponding to the three shortest helices (C, D, and F) could not be determined for lack of available probes. On the basis of the crystal structure of holo-Mb (*7*), helices B to E comprise a compact subdomain. We propose, therefore, that this subdomain unfolds in the transition from the native protein to the I form, while the remainder of the protein (with the possible exception of the F helix) remains folded.

Experimental design. At low temperature and low protein concentration, pH-induced apo-Mb unfolding occurs in two successive stages as monitored by CD (Fig. 1). In a previous study (*6*), aggregation of the apo-Mb I species was observed as a slow irreversible change in the CD signal at certain pH values. To demonstrate that aggregation was not occurring, we measured the concentration dependence of the CD signal and determined that it was linear up to the concentrations used in subsequent experiments (Fig. 2). In addition, the signal at each concentration was stable over time.

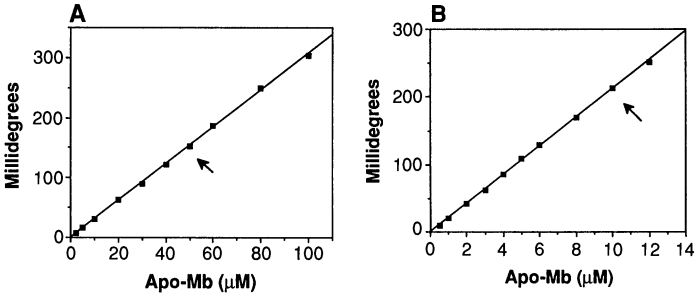

Fig. 2. The concentration dependence of the CD signal at 222 nm (measured in millidegrees) is linear up to the apo-Mb concentrations used in this study (*30*). (**A**) pH 6.0; (**B**) pH 4.2. Arrows mark the concentrations used in exchange-out experiments.

For measuring amide proton exchange rates, apo-Mb in H_2O solution was diluted into buffered D_2O, thereby initiating exchange of amide protons for solvent deuterons. After a period of "exchange-out," further exchange was quenched by the addition of heme and adjusting the pH, thereby reconstituting holo-Mb. Two-dimensional ^1H NMR spectra were acquired after concentrating the reconstituted sample as much as 400 times. Because protons, but not deuterons, are detected by ^1H NMR, the volumes of individual amide proton cross-peaks in COSY spectra serve to monitor the extent of exchange. For a subset of about 50 amide protons, exchange from holo-Mb is very slow (half-times of many days); the other amide protons exchange out during NMR analysis and are unavailable as probes. Assignments have been made for most of the ~50 holo-Mb protons which are stable in D_2O (Fig. 3A) (*8*). By repeating the experiment with various exchange-out periods, the individual rates of exchange can be determined for these protons.

There are two principal advantages of this method. First, exchange-out can be effected under any desired conditions of pH and protein concentration, thereby avoiding conditions under which apo-Mb is known to aggregate and allowing the measurement of exchange rates for both native apo-Mb (at pH 6) and the I form (at pH 4.2). Second, the time resolution of the method is limited only by the time necessary to quench exchange and a slow technique, two-dimensional ^1H NMR, can then be used to assay the extent of exchange at positions throughout the protein. Reconstitution of holo-Mb is rapid in our conditions, as shown by the small extent of exchange-out in the zero timepoint samples (compare Fig. 3, A and B). A kinetic study of the reconstitution of Mb (*9*) shows that both heme binding and protein folding are fast reactions in suitable conditions.

Native and intermediate apo-Mb. Amide proton exchange rates were measured for the native and intermediate forms of apo-Mb. The results are reported as the protection factor $P = k_{exp}/k_{obs}$. The solvent exposed exchange rate k_{exp} is calculated from model peptide data and represents the sequence-, pH-, and temperature-dependent exchange rate expected for an unprotected proton (*10*). The k_{obs} is derived from exchange experiments conducted on native apo-Mb at pH 6.00 (Fig. 3C). For native apo-Mb, P ranges from less than 10 to more than 10^5; for the great majority of the available amide proton probes, P is more than 10^3 (Table 1). As discussed above, only those protons that are stable in reconstituted holo-Mb can be evaluated by this strategy. Nonetheless, our results show that most amide protons that exchange slowly in holo-Mb also do so in native apo-Mb (Fig. 3, A to C, and Table 1). Removing the heme group, therefore, appears to leave the hydrogen-bonded secondary structure intact.

For several protons, P is less than 100; G3 and G4 ($P < 10$) are located near the amino terminus of the G helix in holo-Mb and do not participate in intrahelical hydrogen bonding. Proton H20 ($P = 20$) is surface-accessible in holo-Mb (*11*); exchange rate measurements in many proteins indicate that surface-exposed residues exchange somewhat more rapidly than buried residues. The remaining residues whose amide protons are protected less than 100-fold are CD1 ($P = 30$), E14 ($P = 30$), and G5 ($P = 10$). All of these residues contact the heme group in holo-Mb (*12*). Removal of the heme group may increase local flexibility or solvent penetration (or both) in the heme pocket, leaving these protons relatively unprotected.

Exchange rates for amide protons in the apo-Mb I form were determined by allowing exchange to proceed in D_2O buffered at pH 4.20 (Fig. 3D). Protection factors for amide protons in the I species range from 0.3 to 300 (Table 1). The observation of protection factors less than one is surprising, since it implies that some protons exchange more rapidly than their calculated intrinsic rates. It seems likely that these anomalously low protection factors result from uncertainty in k_{obs}, which in unfavorable cases could be a factor of two, or from factors not taken into account by Molday *et al.* in calculating k_{exp} (*10*). The protection factors observed for the I form

are up to 10,000 times lower than those for native apo-Mb. The relatively modest degree of protection attests to the marginal stability of the I form. These results demonstrate, however, that measurement of amide proton exchange rates is a sensitive tool for mapping regions of structure in a marginally stable state.

In the A, G, and H helix regions, most probe amide protons are at least weakly protected in the I form ($P \geq 5$). These results suggest that the A, G, and H helices are folded in the apo-Mb I form. In contrast, probes in the B and E helix and turn regions all have protection factors of about one, with the exception of B11 ($P = 10$) and B14 ($P = 5$). It appears, therefore, that the B and E helix regions of the apo-Mb I form are in a predominantly unfolded or random coil-like conformation.

Model. Our data led us to suggest a structural model for the apo-Mb native and I forms, based on the structure of holo-Mb. The native state of apo-Mb closely resembles that of holo-Mb. The compactness of native holo- and apo-Mb is indistinguishable as judged by intrinsic viscosity measurements (*4, 13*), which rules out large-scale unfolding on removal of the heme group. Furthermore, the one-dimensional ^{1}H NMR spectrum of native apo-Mb displays substantial chemical shift dispersion, and thus provides evidence that the native state has a distinctive and characteristic conformation (*4*). We have shown (Table 1) that, for most factors of the amide protons that are strongly protected against exchange in holo-Mb, the protection factors are at least 100 in apo-Mb. Therefore, we conclude that, since the pattern of protection observed in holo-Mb is largely preserved in apo-Mb, their structures are similar.

Surprisingly, removal of the heme group leads to a reduction in helix content as judged by CD. Calculations based on the crystal structure of holo-Mb predict a helix content of ~80 percent. The CD measurements of apo-Mb show, however, a helix content of ~55 percent (Fig. 1). Aromatic side chains can affect the CD spectrum in the far ultraviolet (UV) region (*14*), but it seems unlikely that such a large discrepancy can be explained in this way. It is not yet clear what accounts for this apparent reduction in helix content, since virtually all probe protons remain substantially protected (Table 1). Perhaps helices or regions of helices for which no probes are available unfold when the heme group is removed. In particular, the exchange rates for amide protons in the F helix could not be determined because these protons exchange relatively rapidly even in holo-Mb (*15*) and consequently are not trapped by our procedure. Perhaps the F helix is unfolded in native apo-Mb, since in holo-Mb it packs only against the H helix and the heme group (Fig. 4). Another possible explanation for the apparent discrepancy between CD and hydrogen exchange results cannot be excluded; the observed protection may not result from α-helix formation. This appears unlikely because it requires a different conformation that nevertheless produces the pattern of protection found in holo-Mb.

The structure of the apo-Mb I form, by contrast, shows definite unfolding. We propose that the B and E helices, together with the short (seven-residue) C and D helices between them, unfold in passing from native apo-Mb to the I form (Fig. 4). Protection

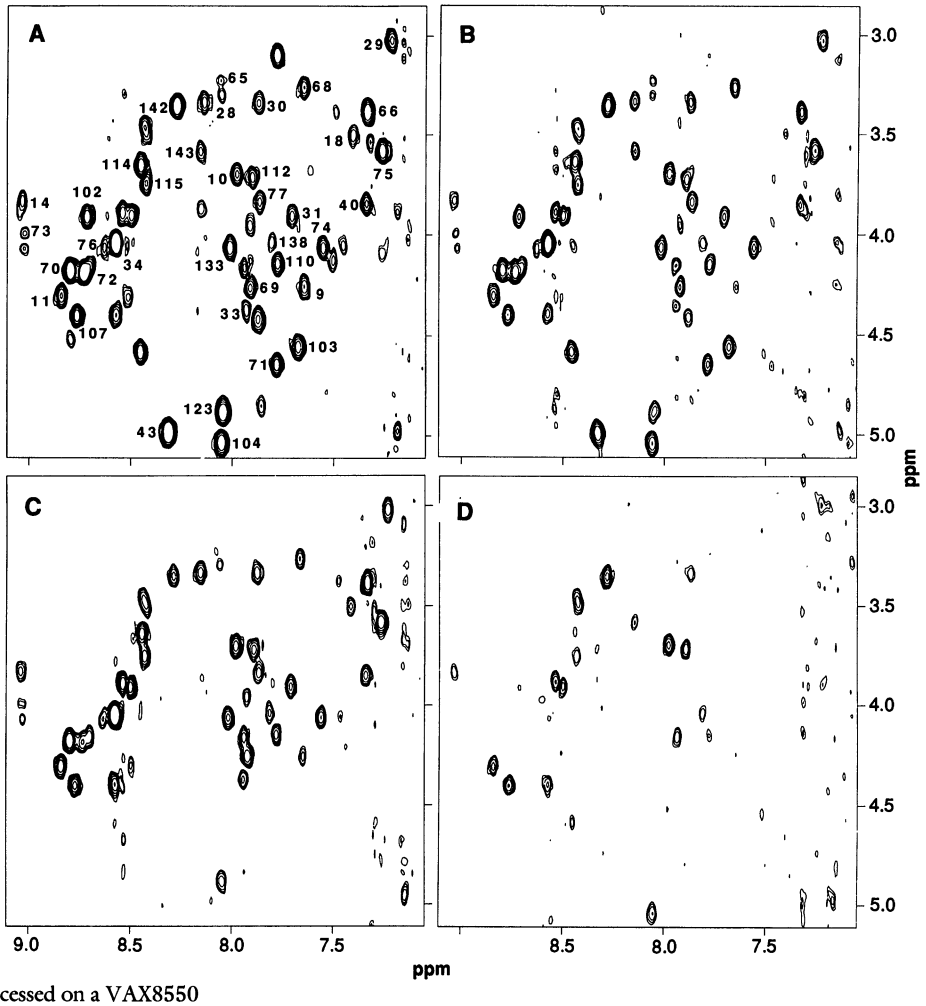

Fig. 3. Two-dimensional spin-correlated (COSY) spectra of Mb-CO in D_2O, pH 5.6, 35°C. (**A**) Native holo-Mb. Myoglobin was dissolved in 2 ml of 10 mM acetate in D_2O, pH 5.6, equilibrated with CO, reduced, and concentrated to 3 mM as described below. Assignments (*8*) used are indicated by residue number. The Val^{17} cross-peak (not shown) lies outside this region at (6.72, 1.58) ppm. (**B** to **D**) Apo-Mb was exchanged out for various times at either pH 6.0 (for native apo-Mb) or pH 4.2 (for the I form) prior to reconstitution. The pH and interval of exchange-out were: (B) pH 6.0, <1 second; (C) pH 6.0, 30 minutes; (D) pH 4.2, 60 minutes. Exchange-out was initiated by diluting an apo-Mb stock solution (>0.5 mM in distilled H_2O) into D_2O buffered with 10 mM acetate. For studies of the native state of apo-Mb, exchange-out conditions were pH 6.00, 50 μM apo-Mb, 5°C. To quench exchange, one-tenth volume of 37 mM DCl and one-tenth volume 0.55 mM bovine hemin, 30 mM NaOD were added in rapid succession. For studies of the I form, exchange-out conditions were pH 4.20, 10 μM apo-Mb, 5°C. To quench exchange, 1/20th volume 0.22 mM hemin in 130 mM NaOD was added. Reconstituted protein (pH 5.6) was concentrated by Amicon filtration (YM-10 membrane) to about 2 ml. Aggregates and residual H_2O were removed by use of a small PD-10 column (Pharmacia) that had been equilibrated with 10 mM acetate in D_2O, pH 5.6. The eluate was equilibrated with CO (15 psi) and reduced with 20 μl of 0.5 M sodium dithionite. The Mb-CO solution was concentrated to 3 to 4 mM as before and incubated for 12 hours at 35°C to allow heme reorientation. In (B), a rapid mixing apparatus (*1*) was used. Magnitude COSY spectra were recorded on a 500-MHz General Electric GN-500 spectrometer. Between 128 and 176 transients of 1K complex data points covering a spectral width of 8264 Hz were recorded for each of 400 t_1 increments (0 to 48 ms). Data were processed on a VAX8550 computer with the program FTNMR, provided by D. Hare, with zero-filling to 2K points and unshifted sinebell multiplication in both dimensions.

Table 1. Amide proton protection factors. Protection factors (k_{exp}/k_{obs}) are given for amide protons for which assignments have been made and for which exchange is effectively quenched by reconstitution. Time points for exchange-out were collected as described in the legend to Fig. 3. The extent of exchange at each time point was determined by calculating the volume integral of each NH–C_αH cross-peak in the COSY spectrum. The data at each time point were normalized with reference to cross-peaks between nonexchangeable protons. Normalized cross-peak intensities for duplicate experiments differed on average by <25 percent. Five time points with exchange-out intervals ranging from <1 second to 3000 minutes were collected for native apo-Mb; seven time points with exchange-out intervals ranging from <1 minute to 630 minutes were collected for apo-Mb in the I (intermediate) form. Nonlinear least squares fitting to an exponentially decaying function were used to determine the observed exchange rates (k_{obs}). Solvent-exposed exchange rates (k_{exp}) were calculated according to (*10*). ND, Not determined. For pH 4.2 data, Leu^{40}, Gly^{65}, and Gly^{73} were lost during reconstitution. For pH 6 data, Leu^{76} could not be measured due to overlap with Lys^{34}.

Residue	Helix position (holo-Mb)	*P* (protection factor)		Residue	Helix position (holo-Mb)	*P* (protection factor)		Residue	Helix position (holo-Mb)	*P* (protection factor)	
		Native (pH 6.0) (apo-Mb)	I (pH 4.2) (apo-Mb)			Native (pH 6.0) (apo-Mb)	I (pH 4.2) (apo-Mb)			Native (pH 6.0) (apo-Mb)	I (pH 4.2) (apo-Mb)
Leu^{9}	A7	9,000	90	Phe^{43}	CD1	30	0.6	Lys^{102}	G3	<10	5
Val^{10}	A8	100,000	200	Gly^{65}	E8	2,000	ND	Tyr^{103}	G4	<10	1
Leu^{11}	A9	>100,000	60	Val^{66}	E9	3,000	1	Leu^{104}	G5	10	8
Trp^{14}	A12	>100,000	20	Val^{68}	E11	1,000	0.4	Ile^{107}	G8	2,000	40
Val^{17}	A15	80,000	40	Leu^{69}	E12	5,000	1	Ala^{110}	G11	6,000	7
Glu^{18}	A16	9,000	10	Thr^{70}	E13	4,000	2	Ile^{112}	G13	60,000	80
Ile^{28}	B9	4,000	2	Ala^{71}	E14	30	0.3	Val^{114}	G15	200,000	30
Leu^{29}	B10	20,000	3	Leu^{72}	E15	100	0.4	Leu^{115}	G16	60,000	7
Ile^{30}	B11	70,000	10	Gly^{73}	E16	3,000	ND	Phe^{123}	GH5	1,000	<1
Arg^{31}	B12	5,000	2	Ala^{74}	E17	1,000	0.8	Lys^{133}	H10	6,000	5
Phe^{33}	B14	6,000	5	Ile^{75}	E18	2,000	1	Phe^{138}	H15	7,000	10
Lys^{34}	B15	4,000	0.9	Leu^{76}	E19	ND	0.6	Ile^{142}	H19	200	60
Leu^{40}	C5	20,000	ND	Lys^{77}	E20	500	<1	Ala^{143}	H20	20	20

observed at B11 and B14 argues for residual structure in the B helix region, but such structure is most likely localized, since little protection is observed at neighboring residues.

In contrast, protection of many amide protons from exchange attests to structure, presumably helical, in the A, G, and H helix regions. In particular, protection is seen at A7, A8, A9, A12, A15, and A16; at G3, G5, G8, G11, G13, G15, and G16; and at H10, H15, H19, and H20 (Table 1). In every case, protected protons are distributed reasonably well throughout the helix region. Among the available probe protons in these helix regions, only G4 fails to show measurable protection, presumably because of its location near the amino terminus of the G helix, as discussed above. A notable feature of this model is that these three helices pack against one another in the holo-Mb structure to form a compact unit (Fig. 4). We propose, therefore, that a compact subdomain consisting of the A, G, and H helices is folded in the apo-Mb I form into a more or less native conformation. We suggest, furthermore, that the B-E helix region is either unfolded into a random coil-like conformation or retains only partial or transient structure (Fig. 4). Whether the F helix participates in the structure seen in the I form is not yet known; its participation would not alter the conclusion that a compact subdomain is folded in the I species. We point out, however, that folding of the A, G, and H helices alone would lead to a helix content of 40 percent, which agrees reasonably well with the measured helix content of the I species (~35 percent).

This model provides a structural rationale for earlier results obtained by site-directed mutagenesis (*6*). Disruptive mutations introduced at G11 strongly destabilized the native state but had a smaller effect on the I form. In the holo-Mb crystal structure, the side chain of G11 forms part of the packing interface between the B and G helices (*12*). Because the native state of apo-Mb closely resembles that of holo-Mb, it is clear that mutations of G11 destabilize the apo-Mb native state by disrupting packing between the B and G helices. It now seems likely that the impact of replacements at G11 on the stability of the I form is muted because the B helix is predominantly unfolded.

Protein folding. We have not established that the stable apo-Mb intermediate studied here is populated during the kinetic process of folding; such a demonstration must await kinetic experiments. Nonetheless, for at least one case in which a stable equilibrium intermediate has been characterized, it appears to correspond to a kinetic intermediate: by several criteria, the A form of α-lactalbumin and an early kinetic intermediate are indistinguishable (*16*). The finding that an apo-Mb subdomain comprising the A, G, and H helices can fold to a structure of moderate stability makes it plausible that this structure constitutes an intermediate on the pathway of apo-Mb folding.

It is also possible to frame the folding problem in a different way. Instead of demanding a kinetic description of the folding process, we seek a hierarchical description of the interactions that lead to stable proteins. Such an approach is implicit in studies of structure formation in short peptides (*17*) and in peptide models of folding intermediates (*18*). A logical extension is to seek residual structure in intact proteins under mildly denaturing conditions. This approach has, however, been frustrated by the cooperativity of protein folding and unfolding although recently it has become clear that it is possible to selectively destabilize the native state, revealing partly unfolded forms. A second difficulty in studying the structure of such species, in addition to their instability, has been their tendency to aggregate, but the proton exchange–NMR technique described above can circumvent this problem by allowing us to work at low protein concentrations.

As discussed above, the observed cooperativity of protein folding leads to the expectation that folding intermediates will be rather unstable. This is true for the apo-Mb I species. Those amide protons that are protected have protection factors of less than 200. The A, G, and H helices each have amide protons with protection factors of 60 or more; thus, complete unfolding of each helix is rather rare. Nonetheless, such protection factors are marginal compared to those seen for native holo-Mb (up to 10^8) (*15*).

Peptides with sequences corresponding to the H helix are substantially helical in aqueous solution (*19, 20*), demonstrating that important determinants of native structure reside at the local secondary structure level and act independently of higher order

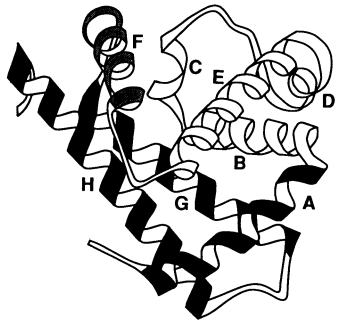

Fig. 4. Ribbon drawing model of Mb (*32*) showing the proposed region of the structured subdomain in the apo-Mb intermediate. Helices B to E are unshaded and are proposed to be predominantly unfolded in the I state. The F helix is lightly shaded to denote the lack of amide proton probes—and consequently, structural information—for this region. The A, G, and H helices are darkly shaded and are proposed to be folded in a native-like conformation in the I form.

interactions. Preliminary CD characterization has failed to reveal helical structure in peptides with sequences corresponding to the G helix (*20*), although a low population of helical conformations has been identified by NMR methods (*21*). Our results suggest that additional helices can assemble onto H. In holo-Mb, A and G each pack against H but do not substantially interact with each other. Whether A and G are marginally stable by themselves and are stabilized by tertiary interactions with H (*5, 12, 22*), or whether they assemble "onto" H by means of an "on-site construction" mechanism (*23*), is not known.

We have not established the detailed nature of tertiary interactions in the I state. The enthalpy of denaturation for the I form is believed to be small (*4*). This observation argues against native-like helix-helix interactions, which should exclude water and result in a substantial denaturation enthalpy. Nonetheless, some type of tertiary interaction (or interactions) must stabilize the A, G, and H helices in the I form, since peptides corresponding in sequence to the G and H helices at most display moderate helix content (above).

The I form of apo-Mb appears to display some of the characteristics of a "molten globule" or "collapsed form" intermediate, which has been observed for some other proteins (*24*). In particular, the I form is rather compact, retains substantial secondary structure, and its hydrophobic side chains are largely exposed to solvent (*4, 25*). The most extensively studied collapsed form is the A form of α-lactalbumin (*3, 26*). This partly folded species is compact and has native-like secondary structure; nonetheless, its side chains are not rigidly locked into position as judged by unfolded-like NMR and near-UV CD spectra (*26*). To determine the structural basis for these observations, Baum *et al.* (*27*) carried out proton exchange experiments similar to those reported here. They found that a number of amide protons are protected from exchange in the A form of α-lactalbumin and that many of these reside in helical portions of the native protein (*27*). Recent amide exchange measurements suggest, however, that other regions of the protein are completely unfolded (*28*). Both for apo-Mb and for α-lactalbumin, therefore, it appears that only a portion of the protein remains folded in the intermediate species.

In the context of this work, it seems reasonable to view folding in terms of the successive assembly of blocks of native-like structure, here termed subdomains. Algorithms have been developed to identify compact subdomains which are contiguous along the polypeptide chain (*29*). Apparently, however, contiguity is not necessary for partial structure: the I form of apo-Mb is folded at both ends of the polypeptide chain, while an intervening subdomain is predominantly unfolded. Similar observations have been made in studies of cytochrome c refolding, in which folding of the NH_2- and COOH-terminal helices is the earliest detectable kinetic event (*2*), and in studies of a peptide model of a one-disulfide intermediate of bovine pancreatic trypsin inhibitor, in which separate peptides form an α-helix and a hairpin loop of β structure when linked by the 30–51 disulfide bond (*18*). It will be of great interest to develop algorithms for predicting stable subdomains a priori.

REFERENCES AND NOTES

1. J. B. Udgaonkar and R. L. Baldwin, *Nature* **335**, 694 (1988).
2. H. Roder, G. A. Elöve, S. W. Englander, *ibid.*, p. 700.
3. K. Kuwajima, K. Nitta, M. Yoneyama, S. Sugai, *J. Mol. Biol.* **106**, 359 (1976); M. Ikeguchi, K. Kuwajima, S. Sugai, *J. Biochem.* **99**, 1191 (1986).
4. Y. V. Griko, P. L. Privalov, S. Y. Venyaminov, V. P. Kutyshenko, *J. Mol. Biol.* **202**, 127 (1988).
5. O. B. Ptitsyn and A. A. Rashin, *Biophys. J.* **3**, 1 (1975); M. Karplus and D. L. Weaver, *Nature* **260**, 404 (1976); D. Bashford, F. E. Cohen, M. Karplus, I. D. Kuntz, D. L. Weaver, *Proteins Struct. Funct. Genet.* **4**, 211 (1988).
6. F. M. Hughson and R. L. Baldwin, *Biochemistry* **28**, 4415 (1989).
7. H. C. Watson, *Prog. Stereochem.* **4**, 299 (1969).
8. T. C. Pochapsky, C. Dalvit, Y. Thériault, P. E. Wright, in preparation.
9. Y. Kawamura-Konishi, H. Kihara, H. Suzuki, *Eur. J. Biochem.* **170**, 589 (1988).
10. R. S. Molday, S. W. Englander, R. G. Kallen, *Biochemistry* **11**, 150 (1972); S. W. Englander and A. Poulsen, *Biopolymers* **7**, 379 (1969).
11. B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971).
12. A. M. Lesk and C. Chothia, *ibid.* **136**, 225 (1980).
13. P. L. Privalov, Y. V. Griko, S. Y. Venyaminov, V. P. Kutyshenko, *ibid.* **190**, 487 (1986).
14. M. C. Manning and R. W. Woody, *Biochemistry* **28**, 8609 (1989).
15. S. S. Narula, Y. Thériault, P. E. Wright, in preparation.
16. M. Ikeguchi, K. Kuwajima, M. Mitani, S. Sugai, *Biochemistry* **25**, 6965 (1986).
17. P. E. Wright, H. J. Dyson, R. A. Lerner, *ibid.* **27**, 7167 (1988).
18. T. G. Oas and P. S. Kim, *Nature* **336**, 42 (1988); J. P. Staley and P. S. Kim, *ibid.* **344**, 685 (1990).
19. J. P. Waltho, V. A. Feher, R. A. Lerner, P. E. Wright, *FEBS Lett.* **250**, 400 (1989).
20. D. Barrick and R. L. Baldwin, unpublished results.
21. J. P. Waltho and P. E. Wright, unpublished results.
22. T. J. Richmond and F. M. Richards, *J. Mol. Biol.* **119**, 537 (1978); M. Gerritsen, K.-C. Chou, G. Némethy, H. A. Scheraga, *Biopolymers* **24**, 1271 (1985).
23. J. Skolnick and A. Kolinski, *Annu. Rev. Phys. Chem.* **40**, 207 (1989).
24. M. Ohgushi and A. Wada, *FEBS Lett.* **164**, 21 (1983); O. B. Ptitsyn, *J. Protein Chem.* **6**, 273 (1987); K. Kuwajima, *Proteins Struct. Funct. Genet.* **6**, 87 (1989); P. S. Kim and R. L. Baldwin, *Annu. Rev. Biochem.* **59**, 631 (1990).
25. P. L. Privalov *et al.*, *J. Mol. Biol.* **205**, 737 (1989).
26. D. A. Dolgikh *et al.*, *Eur. Biophys. J.* **13**, 109 (1985).
27. J. Baum, C. M. Dobson, P. A. Evans, C. Hanley, *Biochemistry* **28**, 7 (1989).
28. J. Baum, C. M. Dobson, P. A. Evans, personal communication.
29. A. M. Lesk and G. D. Rose, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4304 (1981); M. H. Zehfus and G. D. Rose, *Biochemistry* **25**, 5759 (1986).
30. Apo-Mb was prepared from sperm whale myoglobin (Sigma) [F. W. J. Teale, *Biochim. Biophys. Acta* **35**, 543 (1959)]. Samples were buffered with 10 mM (for pH titration or concentration dependence studies) or 1 mM (for spectra) sodium acetate; the pH was lowered beyond 3.4 with HCl. The pH of the sample was measured at room temperature at the conclusion of the experiment. The CD measurements were made at 5.0°C (Aviv 60DS spectropolarimeter) with a 10-mm path length at a protein concentration of 1 μM (determined by UV absorbance in 6 M GuHCl, 20 mM phosphate, pH 6.0 [H. Edelhoch, *Biochemistry* **6**, 1948 (1967)]). For protein concentration dependence studies, 10-mm (pH 4.2) and 1-mm (pH 6.0) path lengths were used.
31. Y.-H. Chen, J. T. Yang, K. H. Chau, *Biochemistry* **13**, 3350 (1974).
32. Schematic ribbon drawing prepared with the aid of the CHAOS program (by D. Richardson and M. Zalis) running entirely within an Evans and Sutherland PS 300 processor. The CHAOS program was provided by N. Tweedy and D. Richardson.

33. We thank D. Lambright, D. Barrick, A. Robertson, Y. Thériault, S. Marqusee, and L. Gavis for advice, assistance, and discussion. Supported by NIH grants GM19988 (R.L.B.) and DK34909 (P.E.W.).

10 May 1990; accepted 24 August 1990