Cooperativity of folding of the apomyoglobin pH 4 intermediate studied by glycine and proline mutations

Yongzhang Luo, Michael S. Kay and Robert L. Baldwin

The apomyoglobin pH 4 folding intermediate contains the A, G, and H helices of myoglobin. Helix destabilizing mutations in the A and G helices are used to test whether the pH 4 folding intermediate of apomyoglobin folds cooperatively. Single glycine or proline mutations destabilize the intermediate substantially, showing that intrinsic helix propensities are important for stability of the intermediate. The A and G helices interact to stabilize each other, as shown by the effect of mutations in the G helix on the unfolding of the A helix, which can be monitored by tryptophan fluorescence. Wild type and the most stable mutant unfold in a two-state reaction, as shown by superposition of the unfolding curves measured by two probes (far-UV circular dichroism and Trp fluorescence), while unfolding of the less stable mutants is more complex. Cooperativity and stability of folding are linked also when stabilizing anions (sulphate, perchlorate) are used to adjust stability.

Sperm whale myoglobin (Mb)[1] contains eight α-helices named A through H. Helices A and G each pack against H but do not substantially interact with each other. The pH 4 folding intermediate (I) of apomyoglobin (apoMb) was found to be formed in a two-state reaction by two different tests. The urea unfolding curves measured by two probes (far-UV circular dichroism (CD) and Trp fluorescence) are superimposable[2] and also the unfolding and refolding kinetics of I fit the two-state model, except for a possible burst phase in unfolding[3]. A novel approach, scanning proline mutagenesis, was used by Schulman and Kim[4] to analyze the cooperativity of folding of an α-lactalbumin folding intermediate: they found that its folding is noncooperative or weakly cooperative. The results and assumptions of these studies are discussed in connection with new results presented here.

Our aim is to test the cooperativity of folding of the pH 4 intermediate by making single glycine or proline mutations at solvent-exposed sites in the A and G helices, which form part of the structure of I[5]. There are two tests. The first is whether the mutations cause a significant loss of stability and, if so, whether the entire folding intermediate or only one helix is destabilized. The second is whether the superposition test for two-state folding is satisfied by the mutant proteins.

An unusual property of Mb makes tryptophan fluorescence and far-UV CD particularly suitable probes. Both Trp residues of Mb, Trp 7 and Trp 14, are in the A helix, and the increased fluorescence of I over U (the unfolded form) results from partial burial of the tryptophan residues. The fluorescence spectrum of I is approximately twice as intense as the spectrum of U, the acid unfolded form, while the wavelength of maximum emission of I is approximately halfway between U and N (M. Jamin & R.L.B., submitted). These properties indicate that the increased fluorescence of I over U results from partial burial of the two Trp residues away from solvent, and burial must be dependent on formation of the A helix. The structure in which the two Trp residues are partly buried is probably the A, G, H interface, as in the structure of myoglobin[1], but alternative structures are possible. Far-UV CD (222 nm) measures the overall helix content of I, which has a partly helical, partly unfolded CD spectrum[5].

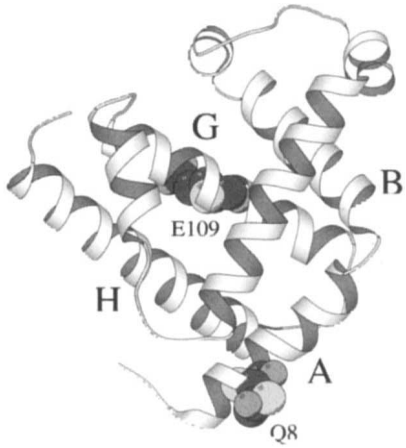

Fig. 1 Molscript[24] diagram of sperm whale myoglobin showing the locations of the mutations.

A similar superposition test, comparison of helix content monitored by far-UV CD with side chain burial monitored by fluorescence or absorbance, has been used successfully to detect folding intermediates in the unfolding transitions of those few native proteins showing intermediates: for example, carbonic anhydrase[6] and α-lactalbumin[7]. These studies suggest that the superposition test used

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305-5307, USA.

Correspondence should be addressed to R.L.B. *email: yluo@cmgm.stanford.edu*

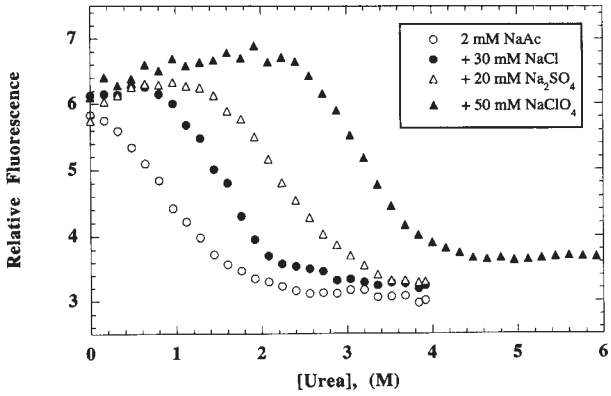

Fig. 2 Fluorescence-monitored urea unfolding of WT at pH 4.2, 4 °C, in four sets of anion conditions: 2 mM sodium acetate, 30 mM NaCl, 20 mM Na_2SO_4, and 50 mM $NaClO_4$. The buffer is 2 mM sodium acetate when salt is present.

here and earlier[2] should be a sensitive test for intermediates in the folding of I.

The pH 4 intermediate of apoMb provides an interesting test system for investigating intermediates formed in conditions where the native protein refolds, for the following reasons. (i) The pH 4 folding intermediate contains just three helices (A, G, H) of the eight helices of Mb, as shown by NMR-hydrogen exchange[5]. Since the A and H helices come together from opposite ends of the polypeptide chain, and since the A, G, and H helices form a compact subdomain in native Mb, this result suggests that I has specific structure that resembles a subdomain of Mb. (ii) The pH 4 folding intermediate is also a transient intermediate in the refolding kinetics of native apoMb: I is formed rapidly (within a few ms) when native apoMb refolds at pH 6[8]. Thus, studies of the properties of I at equilibrium (pH 4) provide information about an important transient intermediate in the refolding of native apoMb. (iii) The B helix is induced to become a stable part of I, in addition to the A, G, and H helices, by adding a stabilizing anion (20 mM trichloroacetate)[9]. Thus, the structure of I is not fixed but instead depends on environmental conditions; probably the stability of I is a main factor. It may seem paradoxical that I should have specific structure that nevertheless may change with conditions, but these two propositions are not necessarily in conflict.

Design and properties of the mutants

Our plan was to make a small number of mutants and characterize them thoroughly. We chose two solvent-exposed sites (Fig. 1), one in the A helix (Gln 8) and the other in the G helix (Glu 109), and substituted Gly and Pro at each site. In the native Mb[1], the A helix includes residues 3–18 and the G helix includes residues 100–118. Our expression system yields native proteins (in the case of these mutants, a mixture of native apo and holoMb is obtained), and the expression yield depends on the stability of the native form. We were unable to express Q8P, and the yields of Q8G, E109G, and E109P were only 20%, 10%, and 5% respectively, of the yield for WT. Glycine was chosen as a helix-destabilizing residue, in addition to proline, because proline in a protein helix frequently bends the helix and a bent helix may be sufficient to cause noncooperative folding. After proline, glycine is the most helix-destabilizing residue in alanine-based peptides[10].

The chosen substitution sites are solvent-exposed so that we can examine the effects of helix-destabilizing mutations without affecting possible packing interactions between helices. The substitution sites are well within each helix in order that the mutations have a strong helix-destabilizing effect, and the original residues at the chosen sites are good helix formers for the same reason. Also, the sites chosen are not conserved in the globin family[11].

The pH unfolding curves of the mutants measured by CD show normal U $\rightleftharpoons$ I unfolding transitions near pH 3 and I is fully populated at pH 4 in each mutant (data are not shown: CD-monitored pH unfolding curves were measured in 2 mM Na citrate; data on the helix contents of the pH 4 intermediates are given in Table 2). The native form tends to aggregate in the case of the mutants and their I $\rightleftharpoons$ N transitions are not well resolved in pH unfolding curves. The fluorescence emission spectra of the mutant intermediates have a similar shape as WT (maximum at ~340 nm, measured in 5 mM Na acetate, pH 4.2), but the fluorescence intensity is altered, and the less

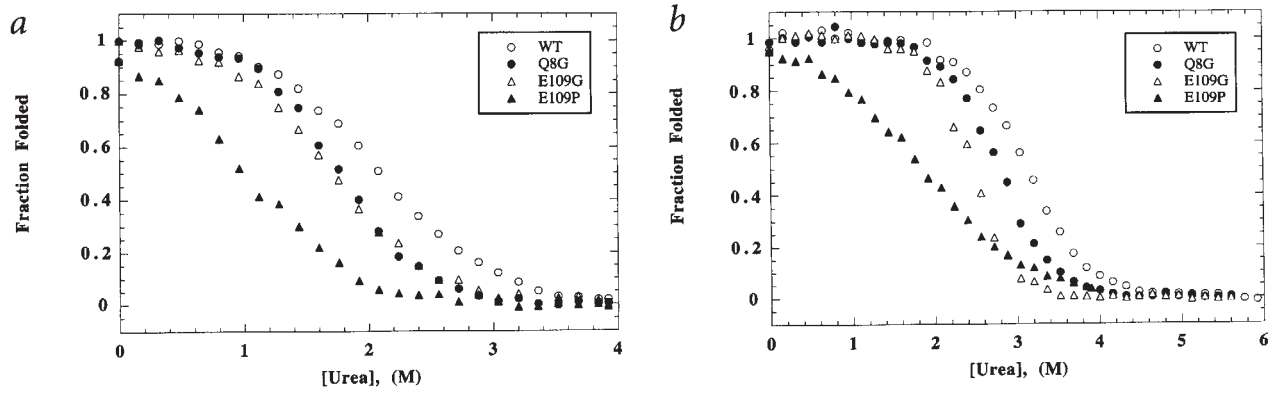

Fig. 3 Fluorescence-monitored unfolding (normalized data) of WT and the three mutants in ***a***, 20 mM Na_2SO_4, and ***b***, $NaClO_4$.

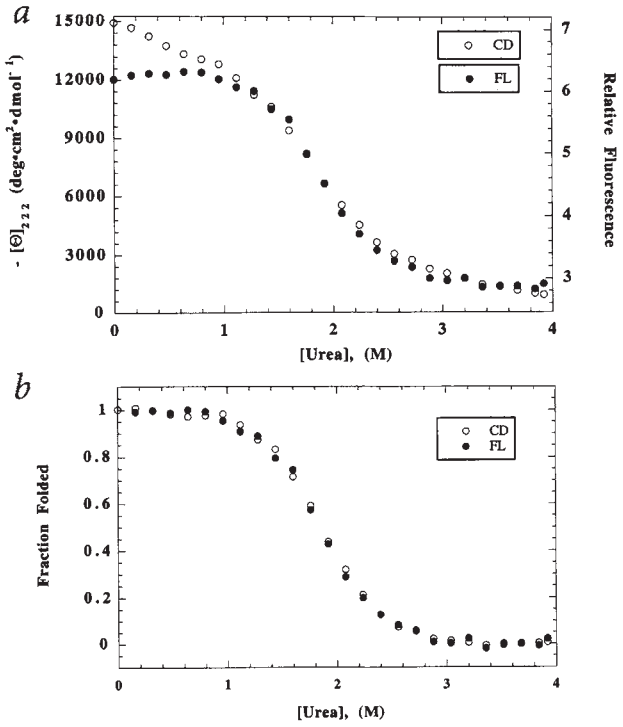

Fig. 4 Test for superposition of the CD- and fluorescence-monitored unfolding curves of Q8G in 20 mM Na_2SO_4, 2 mM NaAc, pH 4.2, 4 °C. ***a***, Raw data; ***b***, normalized. The fluorescence data point below 0.16 M urea is omitted because of a curving baseline: see data analysis in the Methods section for explanation.

stable mutants E109G and E109P have much reduced intensities. Spectra taken in the presence of a stabilizing anion (20 mM Na sulphate) are presented in Fig. 6.

Mutations and anions affect stability

The fluorescence-monitored unfolding curves of I are shown in Fig. 2, in four different sets of anion conditions. Goto *et al.*[12] showed that various anions induce the folding of the acid forms of apoMb and cytochrome *c* at pH 2, and the anions differ considerably in the concentrations needed to induce folding. The data in Fig. 2 show that the anions also have characteristic and widely different effects on the stability of the pH 4 intermediate; Loh *et al.*[9] observed earlier that 20 mM trichloroacetate has a remarkable stabilizing effect. Trichloroacetate is not used here because it has a severe photobleaching effect on Trp residues. Perchlorate also has a photobleaching effect, but it is relatively mild. The ability to adjust stability by adding an anion is important for our study, because the urea unfolding curves of the less stable mutants are cut off at 0 M urea when only weakly stabilizing conditions are used, such as 30 mM NaCl.

Fig. 3 shows the fluorescence-monitored unfolding curves of WT and the three mutants in 20 mM Na_2SO_4 (Fig. 3*a*) and in 50 mM $NaClO_4$ (Fig. 3*b*). The curves are normalized by subtracting the baselines, by using the procedure of Santoro and Bolen[13]. The order of relative stability is WT > Q8G > E109G > E109P, when relative stability is measured by C_m, the midpoint of the urea unfolding transition. A study by Fersht and coworkers[14] concludes that C_m values provide the best measure of relative stability in a series of mutants, all of which have approximately the same m-values, because the use of C_m avoids the long extrapolation of ΔG values to 0 M urea to obtain ΔG^{H_2O}. When the linear extrapolation model ($\Delta G = \Delta G^{H_2O} - mC$) is used, then $\Delta G^{H_2O} = mC_m$ and the rank order of C_m values is the same as that of ΔG^{H_2O}. For mutants whose apparent m-values are significantly different from the WT value, the change in m-value is likely to reflect a deviation from two-state folding behaviour, and the rank order of C_m values for the mutants still provides a sensible measure of relative stability in this case. The C_m values from Fig. 3, and also those found from CD-monitored unfolding, are given in Table 1.

The main conclusion from Fig. 3 is that mutations in the G helix, E109G and E109P, lower the stability of the entire folding intermediate including the A helix, when unfolding is monitored by Trp fluorescence. By this test, the A and G helices are stabilized in I by interacting with each other, possibly through the H helix, and by this test the folding of I is cooperative.

Table 1 shows that the same C_m values are obtained, within error, for unfolding monitored by CD and by fluorescence for WT and also for the most stable mutant, Q8G. As the mutant stability decreases, it becomes progressively more difficult to obtain values of C_m from CD-monitored unfolding. Table I also shows that the order of stability provided by the different anion conditions (50 mM $NaClO_4$ > 20 mM Na_2SO_4 > 30 mM NaCl) is the same for WT and each of the three mutants, except that values of C_m are not obtainable for E109G and E109P in 30 mM NaCl.

Superposition test for two-state folding

The superposition test for two-state folding applied to the most stable mutant, Q8G, in 20 mM Na_2SO_4, is shown in Fig. 4. The raw data for unfolding monitored by CD and fluorescence

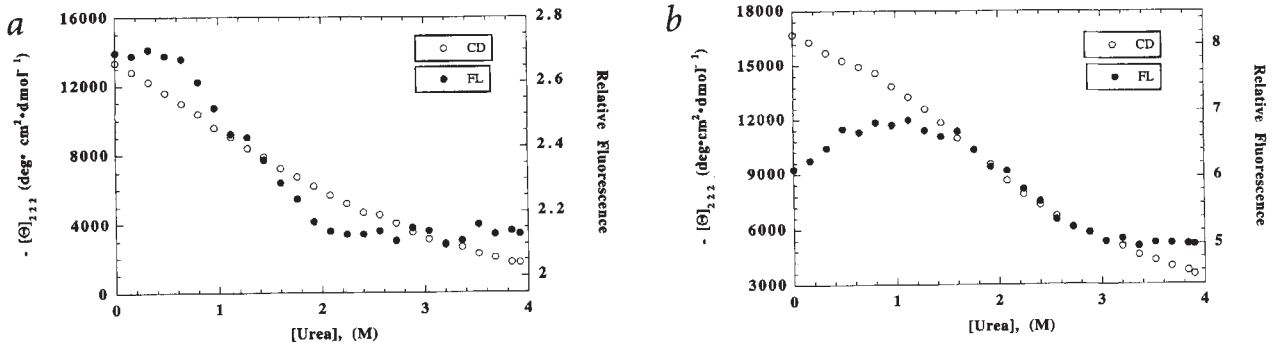

Fig. 5 The superposition test for cooperative folding of E109P: ***a***, in 20 mM Na_2SO_4; ***b***, in 50 mM $NaClO_4$.

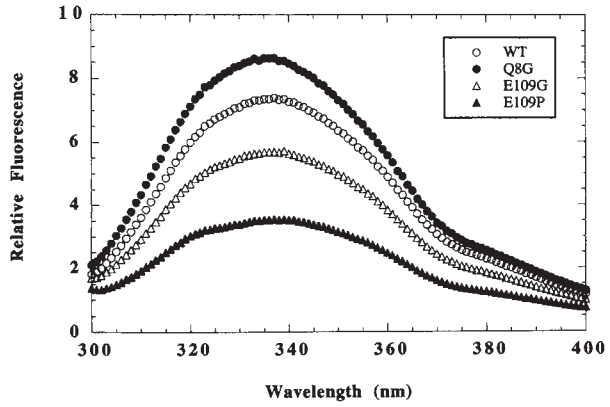

Fig. 6 Fluorescence emission spectra of WT and the three mutants in 20 mM Na_2SO_4.

are shown in Fig. 4*a* and the normalized data, after subtraction of the baselines, are shown in Fig. 4*b*. The two normalized curves agree within the error of finding the baselines and the Q8G mutant folds in a two-state reaction, as reported earlier for WT[2]. This experiment provides direct evidence that a mutation in the A helix destabilizes the entire folding intermediate, because the unfolding curves monitored by fluorescence and CD shift simultaneously to lower urea concentrations. The same conclusion is reached in a different way from the data in Fig. 3, which shows that either a Gly or Pro substitution for E109 in the G helix destabilizes the A helix, as monitored by the fluorescence change originating from the two Trp residues in the A helix.

The superposition test applied to the least stable mutant, E109P, is shown in Fig. 5. The CD-monitored unfolding curve is quite broad, almost linear, and it shows no clear evidence of an unfolding transition. It is not possible to define a meaningful transition curve by subtracting assumed baselines. On the other hand, CD-monitored unfolding of E109P in 50 mM $NaClO_4$ (Fig. 5*b*), in which the stability is increased, shows some indication of an unfolding transition and partial juxtaposition of the unfolding curves found by CD and fluorescence can be achieved.

The main conclusion from Fig. 5*a* is that the two-state folding found for WT and Q8G breaks down in the case of the least stable mutant E109P in 20 mM Na_2SO_4. The behaviour of E109G is intermediate between that of E109P and Q8G. On the other hand, a partial regain of cooperative folding is found for E109P in more stable conditions, in 50 mM $NaClO_4$ (Fig. 5*b*). There is a qualitative correlation between the stability and cooperativity of folding when stabilizing anions as well as mutations are used to alter stability (Table 3), and C_m is used to estimate stability (Table 1).

Table 1 Stability[1] to urea unfolding of apomyoglobin mutants

ApoMb	30 mM NaCl	20 mM Na_2SO_4	50 mM $NaClO_4$
WT-FL[2]	1.42	2.08	3.14
WT-CD[3]	1.45	2.09	3.24
Q8G-FL	0.80	1.85	2.84
Q8G-CD	0.75	1.86	2.84
E109G-FL	NM[4]	1.73	2.64
E109P-FL	NM	1.01	1.94

[1]Stability is given as C_m, the urea molarity at the transition midpoint. The conditions are 2 mM NaAc with different anions, as indicated, at pH 4.2, 4 °C.
[2]FL: Tryptophan fluorescence emission at 330 nm (with excitation at 280 nm).
[3]CD: circular dichroism at 222 nm.
[4]NM: not measurable.

Comparison with the results for α-lactalbumin

Cooperative folding is one of the important properties of small, single-domain proteins. It is not yet known at what stage in the folding process cooperative folding first emerges. Our approach to the problem is to characterize the folding cooperativity of the pH 4 equilibrium intermediate, which also occurs as a transient intermediate in the refolding kinetics of the native protein[8]. Several approaches to the problem of measuring cooperativity in equilibrium folding intermediates have been tried. The problem is exacerbated by the fact that most folding intermediates aggregate, except at extremely low protein concentrations. NMR approaches are made difficult both by aggregation and by poor chemical shift resolution, because the side chain packing which gives highly resolved chemical shifts in native proteins is typically not present in folding intermediates. The calorimetric criterion, based on comparing van't Hoff and calorimetric enthalpies of unfolding, remains controversial when applied to folding intermediates (compare Griko and Privalov[15]; and Nishi *et al.*[16]), in part because the apoMb intermediate, at least, shows both heat and cold denaturation[16].

The approach introduced by Schulman and Kim[4] has produced very interesting results for an α-lactalbumin folding intermediate. They made single and double proline mutations in individual helices and showed that the overall helix content measured by CD dropped by approximately the amount expected if the proline destabilizes only the helix where the mutation is made. They were unable to measure urea or GdmCl unfolding curves for technical reasons (P.S. Kim, personal communication) and so any changes in stability that may be obtained from the unfolding curves remain unknown. In contrast, the proline and glycine mutations made here affect the stability of the entire folding intermediate, as measured by the urea midpoint values in Table 1, but they do not affect the overall helix content measured by CD (Table 2). Thus, the α-lactalbumin and apomyoglobin intermediates respond very differently to helix-destabilizing mutations, and the folding behaviours of these two intermediates evidently differ in a fundamental way.

On the other hand, the highly cooperative folding observed here for the AGH subdomain in WT apoMb and the Q8G mutant breaks down for the least stable mutant E109P in 20 mM Na_2SO_4, as shown by the superposition test (Fig. 5*a*). It is quite possible that the α-lactalbumin intermediate resembles E109P in its folding cooperativity. We are unable to test this hypothesis by making double proline mutants in one helix, as

Table 2 Helix content[1] of the pH 4 intermediate at 0 M urea in different anion conditions

ApoMb	5 mM NaAc	30 mM NaCl	20 mM Na_2SO_4	50 mM $NaClO_4$
WT	14,100	14,300	14,700	16,800
Q8G	13,200	14,500	14,900	17,500
E109G	13,200	13,600	14,700	17,800
E109P	12,200	12,700	13,300	16,700

[1]The buffer is 2 mM NaAc, at pH 4.2, 4 °C, except when only NaAc is present. Helix content is measured by circular dichroism at 222 nm, and is given as the negative mean residue ellipticity (deg•cm^2•dmol^{-1}).

Table 3 Dependence of cooperativity of folding[1] on intrinsic stability and anion conditions

ApoMb	5 mM NaAc	30 mM NaCl	20 mM Na_2SO_4	50 mM $NaClO_4$
WT	+[2]	+	+	+
Q8G	-	+	+	+
E109G	-	-	±	±
E109P	-	-	-	±

[1]Cooperativity (apparent 2-state behavior) was tested by superposition of the fluorescence- and CD-monitored unfolding curves, each of which was fitted to a two-state equation ($U \rightleftharpoons I$). The stability of apo I increases with the different anions from left to right, and with the various mutants from bottom to top (see Table 1).
[2]"+" means cooperative; "-" means not cooperative; "±" means nearly cooperative.

Schulman and Kim did[4], because our expression system requires that the native protein be moderately stable. Our results with different anion conditions confirm the conclusion from the mutant studies that cooperativity is closely linked to the stability of folding (Table 3; see also Fig. 5*a*,*b*).

The contrast between the broad, almost linear, CD unfolding curves and the clearly sigmoidal fluorescence unfolding curves of E109P in 20 mM Na_2SO_4 (Fig. 5*a*) is surprising and puzzling. The change in fluorescence intensity upon unfolding is quite small in E109P compared to WT. The complete spectra are given in Fig. 6. The shape of the spectrum is similar for WT and the three mutants, but the intensity is much reduced in E109P and E109G. If the increase in intensity of I over U is caused by partial burial of the two Trp residues in I, then the spectra in Fig. 6 mean that the extent of burial in E109P is much less than in WT or Q8G. This can happen either if only a small fraction of the intermediate population has buried Trp residues in E109P, or if the whole population has Trp residues buried only to a slight extent. The contrast between the CD unfolding curve, which suggests a broad distribution of intermediate species with varying stabilities to urea unfolding, and the fluorescence unfolding curve, which suggests a more homogeneous intermediate, can be used to argue that the first explanation is more likely.

Stabilization of early folding intermediates

How are helices stabilized in early folding intermediates? There is general agreement that the hydrophobic interaction is chiefly responsible, because the side chains of folding intermediates are not locked in place and strong specific side chain interactions are unlikely. Calculations by Richmond and Richards[17] show that the hydrophobic interaction is long-range: when two helices of myoglobin are moved towards each other in fixed orientation, the total accessible surface area begins to decrease when the two helices are still 6 Å apart. Because the helices of native proteins usually have one hydrophobic face and one hydrophilic face, a long-range hydrophobic interaction can stabilize native helices when they are packed together quite loosely[18]. This loose hydrophobic interaction may account for the very weak cooperativity of helix stabilization observed by Schulman and Kim[4] in a folding intermediate of α-lactalbumin. Isolated peptides from helix-containing regions of proteins rarely show stable helix formation in water[19], and therefore interactions of some kind with the rest of the protein must stabilize a given helix.

We observe here that cooperativity of folding is directly linked to stability for the pH 4 folding intermediate of apoMb. An inverse relation had been expected: stability should increase when folding becomes cooperative, because cooperative folding probably involves partial close packing and thus strengthening of the hydrophobic interactions. We propose that the kinetic folding reaction, which is not cooperative at the start of the folding process, switches to a cooperative mode as the intermediate acquires sufficient stability.

Methods

Construction of mutants and protein preparation. Mutants Q8G, Q8P, E109G, and E109P were constructed from the synthetic sperm whale myoglobin gene from pMbT7 plasmid[9] using the polymerase chain reaction. DNA sequences were verified by standard dideoxy methods in the PAN facility of Stanford University. All molecular biology techniques were performed according to Sambrook *et al.*[20] with minor modifications. Myoglobin was expressed in *E. coli* strain BL21/DE3 (Novagen). Hemin (Sigma) dicyanide (Mallinckrodt) was added to the lysozyme-lysed cells to convert any apomyoglobin to holomyoglobin, which was purified to more than 95% homogeneity based on SDS-PAGE as described[9]. The haem group was removed by the acid methylethylketone (MEK, J.T. Baker) method[21], which requires care to remove traces of MEK but gives a better yield than the acid-acetone method[22]. Apomyoglobin concentration was determined in 6 M guanidine HCl (Gibco BRL), pH 6.0, by UV absorbance at both 280 nm and 288 nm, assuming ε_{280} = 15, 200 M^{-1} cm^{-1} and ε_{288} = 10, 800 M^{-1} cm^{-1} (ref. 23).

CD and fluorescence measurements. The CD-monitored helical content at 222 nm did not vary with protein concentrations between 0.5 to 50 μM for all the mutants at pH 4.2 (data not shown). The buffer used for measuring the urea (USB)-induced unfolding curve is 2 mM sodium acetate (J.T. Baker) when another stabilizing anion is present, or 5 mM sodium acetate in Table 2. The mutants showed a single peak in FPLC superose-12 column (Pharmacia Biotech) analysis under all of the unfolding buffer conditions. Unfolding is reversible as tested by refolding the unfolded protein in 6 M urea, after dilution into buffer at pH 4.2, 4 °C, or by unfolding apoMb I initially present in 0 M urea. CD data were collected at 222 nm with a 1 cm cuvette, using an Aviv 62A DS circular dichroism spectropolarimeter. Fluorescence measurements were made using a SLM AB2 spectrofluorimeter with a 1 cm light path. Excitation was at 280 nm (with a bandpass of 4 nm) and emission was monitored at 330 nm (with a bandpass of 16 nm). Protein concentrations of 1 μM and 2 μM were used for fluorescence and CD measurements respectively.

Data analyses. The urea-induced unfolding curves were analyzed by assuming a two-state model with linear baselines, according to the procedure of Santoro and Bolen[13], which uses data inside as well as outside the transition zone to fix the baselines. The fluorescence data points between 0 and 1 M urea were omitted from the analysis if they show the fluorescence increasing with urea concentration. Work by M. Jamin and R.L.B. (submitted) shows that the rising fluorescence results from an equilibrium between two forms ($I_a \rightleftharpoons I_b$) in which I_a has the higher fluorescence. Data for the unfolding transition (above 1 M urea) pertain to the urea unfolding of I_a.

Acknowledgements
We thank S. Loh for his suggestions on protein preparation and W. Fiori for his assistance with the Molscript diagram of sperm whale myoglobin. Y. Luo is a fellow of the National Institute of General Medical Sciences. M.S. Kay is a trainee of the National Institute of General Medical Sciences Medical Scientist Training program. This work was supported by a National Institutes of Health Grant to R.L.B.

Received 23 April, 1997; accepted 16 September, 1997.

1. Takano, T. Structure of myoglobin refined at 2.0 Å resolution I. crystallographic refinement of metmyoglobin from sperm whale. *J. Mol. Biol.* **110**, 537–568 (1977).
2. Kay, M.S. & Baldwin, R.L. Packing interactions in the apomyoglobin folding intermediate. *Nature Struct. Biol.* **3**, 439–445 (1996).
3. Jamin, M. & Baldwin, R.L. Refolding and unfolding kinetics of the equilibrium folding intermediate of apomyoglobin. *Nature Struct. Biol.* **3**, 613–618 (1996).
4. Schulman, B.A. & Kim, P.S. Proline scanning mutagenesis of a molten globule reveals noncooperative formation of a protein's overall topology. *Nature Struct. Biol.* **3**, 682–687 (1996).
5. Hughson, F.M., Wright, P.E., & Baldwin, R.L. Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548 (1990).
6. Wong, K.P. & Tanford, C. Denaturation of bovine carbonic anhydrase B by guanidine hydrochloride. A process involving separable sequential conformational transitions. *J. Biol. Chem.* **248**, 8518–8523 (1973).
7. Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. Three-state denaturation of α-lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* **106**, 359–373 (1976).
8. Jennings, P.A. & Wright, P.E. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **262**, 892–896 (1993).
9. Loh, S.N., Kay, M.S. & Baldwin, R.L. Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl. Acad. Sci. USA* **92**, 5446–5450 (1995).
10. Rohl, C.A., Chakrabartty, A. & Baldwin, R.L. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.* **5**, 2623–2637 (1996).
11. Lesk, A.M. & Chothia, C. How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. *J. Mol. Biol.* **136**, 225–270 (1980).
12. Goto, Y., Takahashi, N. & Fink, A.L. Mechanism of acid-induced folding of proteins. *Biochemistry* **29**, 3480–3488 (1990).
13. Santoro, M.M. & Bolen, D.W. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988).
14. Serrano, L., Kellis, J.T. Jr, Cann, P., Matouschek, A. & Fersht, A.R. The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**, 783–804 (see page 785) (1992).
15. Griko YV; Privalov PL. Thermodynamic puzzle of apomyoglobin unfolding. *J. Mol. Biol.* **235**, 1318–1325 (1994).
16. Nishii, I., Kataoka, M. & Goto, Y. Thermodynamic stability of the molten globule states of apomyoglobin. *J. Mol. Biol.* **250**, 223–238 (1995).
17. Richmond, T.J. & Richards, F.M. Packing of alpha-helices: geometrical constraints and contact areas. *J. Mol. Biol.* **119**, 537–555 (1978).
18. Baldwin, R.L. How does protein folding get started? *Trends Biochem. Sci.* **14**, 291–294 (1989).
19. Muñoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct. Biol.* **1**, 399–409 (1994).
20. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular Cloning: a laboratory manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).
21. Teale, F.W.J. Cleavage of the haem-protein link by acid methylethylketone. *Biochim. Biophys. Acta.* **35**, 543–543 (1959).
22. Fanelli, A.R., Antonini, E. & Caputo, A. Studies on the structure of hemoglobin. I. Physicochemical properties of human globin. *Biochim. Biophys. Acta.* **30**, 608–615 (1958).
23. Edelhoch, H. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**, 1948–1954 (1967).
24. Kraulis, P.J. Molscript - A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).