Packing interactions in the apomyglobin folding intermediate

Michael S. Kay and Robert L. Baldwin

The contribution of specific packing to the stability of the sperm whale apomyoglobin intermediate has been studied by urea denaturation monitored by circular dichroism and fluorescence. Mutations disrupting native packing sites within the subdomain formed by the A, G and H helices destabilize the intermediate, in contrast to the conclusion drawn from earlier studies of pH-induced unfolding. Based on these results, the intermediate is proposed to be stabilized by both partially formed native-like tertiary, and non-specific hydrophobic interactions forming a subdomain folding intermediate. The results help to explain how the intermediate acquires its structure and stability.

Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305-5307, USA

Correspondence should be addressed to R.L.B.

Protein folding intermediates typically have some secondary structures found in the native conformation, but the amount of tertiary structure and the extent to which it resembles the native tertiary structure, remain unsolved mysteries. This problem can be studied best with folding intermediates that are stable at equilibrium, such as the pH 4.2 (I_1) intermediate of apomyoglobin (apoMb). Our laboratory reported previously that the pH-induced unfolding transition curves of I_1 are relatively insensitive to mutations that introduce packing defects[1]. We report here that the urea-induced unfolding transition curve is, however, sensitive to packing mutations, and we use this behaviour to analyse how closely I_1 resembles N, the native conformation of apoMb.

Sperm whale myoglobin is a small (M_r 17,000), α-helical protein that has served as a protein folding model system for over 30 years[2]. It is monomeric and has eight α-helices labelled A–H (Fig. 1). On removal of haem, apoMb remains a well-folded protein with extensive secondary and tertiary structure at neutral pH (the native state, N)[3–5]. Under mildly acidic conditions (pH 4.2), apoMb forms a partially folded intermediate (I_1 or I) with physical properties intermediate between the native and unfolded states[6–8]. Similar intermediates have been studied at pH 2 or 3 in the presence of high salt (>250 mM NaCl) [8]. A more structured intermediate, I_2, is formed in the presence of 20 mM trichloroacetic acid at pH 4.2 and has been recently characterized by hydrogen exchange[9]. I_1 is compact, has significant secondary structure, and lacks fixed tertiary structure as measured by near-UV CD[3,7,10] and 1-D NMR[9]. Hydrogen exchange studies[11] on this intermediate detected significant protection only in the A, G and H helices. Since these three helices form a compact subdomain in the native myoglobin structure (Fig. 1), the intermediate was proposed to be comprised solely of this subdomain, with the rest of the protein largely unfolded. A kinetic intermediate formed in milliseconds has a similar pattern of protection[12].

An immediate question raised by the AGH subdomain model is the source of the intermediate's stability. In the native structure, the AH and GH helical interfaces comprise two of the largest hydrophobic helix contacts in the protein[13,14]. The A, G, and H helices are strongly amphipathic, suggesting that hydrophobic packing between helices is important in defining the myoglobin architecture. Studies of isolated peptides corresponding to the A, G, and H helices revealed little or no helical structure[1,2,15], ruling out intrinsic helical propensities as the dominant stabilizing force. Peptides corresponding to the isolated G and H helices show no stabilizing interaction in mixing experiments at concentrations up to 1 mM[1].

The importance of native-like hydrophobic packing interactions was explored by Hughson *et al.*[1] using mutagenesis of key helix packing sites among the A, G and H helices in the native structure. These mutations had a profound effect on the stability of the native state of apomyoglobin, increasing the pH midpoint for the N→I transition up to 0.8 pH units and decreasing the C_m for urea denaturation (N→U) by as much as 1.2 M urea. Surprisingly, these mutations generally had only slight effects on the pH midpoint of the intermediate to unfolded transition, leading to the conclusion that native-like packing is not present in the intermediate which therefore can be aptly termed a molten globule intermediate. A loose, nonspecific hydrophobic interaction that stabilizes the three amphipathic helices was

proposed. Such a diffuse and fluctuating interface would be expected to resist mutagenesis. Interestingly, the addition of larger hydrophobic side chains in these helix packing sites slightly increased the stability of the intermediate (that is, decreased the pH midpoint), supporting the loose interaction model and the characterization of this intermediate as a molten globule species.

In this study, we re-examine the role of native-like packing interactions in the stability of the apomyoglobin folding intermediate. Instead of utilizing CD-detected pH denaturation curves to measure stability changes, we employ urea denaturation measured by both CD and fluorescence. Contrary to previous observations, all of the native hydrophobic helix packing sites examined are found to contribute significantly to the stability of the intermediate. The observation of native-like packing helps to explain how the intermediate achieves its structure and stability in the absence of significant helical propensity of its constituent helices.

Analysis of intermediate stability

At pH 4.2, the I form of apomyoglobin is >95% populated[16]. Examination of urea-induced unfolding at pH 4.2 allows quantitative measurement of the intermediate stability. Urea-induced unfolding curves for I were measured at pH 4.2, 4 °C in 4 mM citrate buffer. These curves were fitted to a two-state equation with linear baselines for the folded and unfolded states: the baselines are determined using data inside as well as outside the transition zone[17]. The large slope of the native baseline for both CD and fluorescence is a major source of error in fitting the transition, but the use of both CD and fluorescence data reduces this error significantly. The nature of these baseline effects is not known but is of current interest as it may provide a clue to understanding deviations from precise two-state behaviour[8]. CD and fluorescence data are nearly superimposable after normalization (Fig. 2), which subtracts the strong baseline contributions from the urea unfolding data. Normalization allows a direct comparison of CD and fluorescence data between mutants (Figs 3, 4).

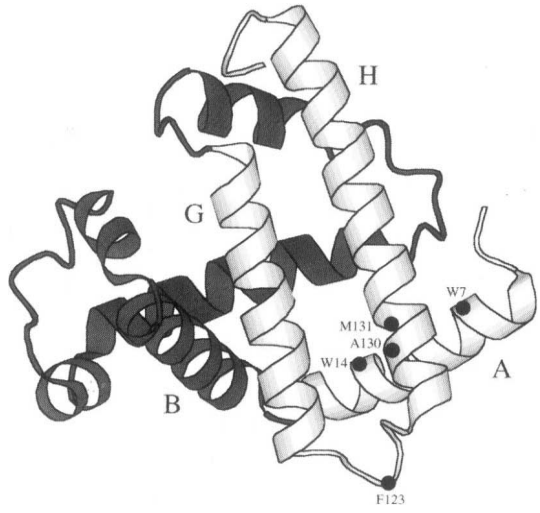

Fig. 1 Molscript[39] diagram of sperm whale myoglobin with AGH subdomain highlighted (white). Mutated residues within the AGH subdomain are highlighted.

In fitting the data to a two-state transition, the dependence of ΔG on denaturant concentration (m-value) is generally poorly determined. The m-value should not change among mutants since the mutations made here are not expected to have a large effect on the surface area either of the unfolded form or of the folding intermediate[18]. Fixing the m-value prevents random error in determining m from affecting the measurement of the apparent free energy. Essentially, this method gives a quantitative measure of the midpoint of urea denaturation (C_m), transforming the C_m values into apparent free energies (ΔG_{app}). To facilitate comparison of mutants, m-values were fixed at the wild-type value, as measured by CD. Fluorescence-monitored urea unfolding curves consistently have slightly higher m-values than their CD counterparts (also reported by Eftink and co-workers[19]). This difference may be caused by residual ellipticity measured by CD after the tryptophans are fully exposed. Fixing the m-value at the CD-derived value yields similar sets of

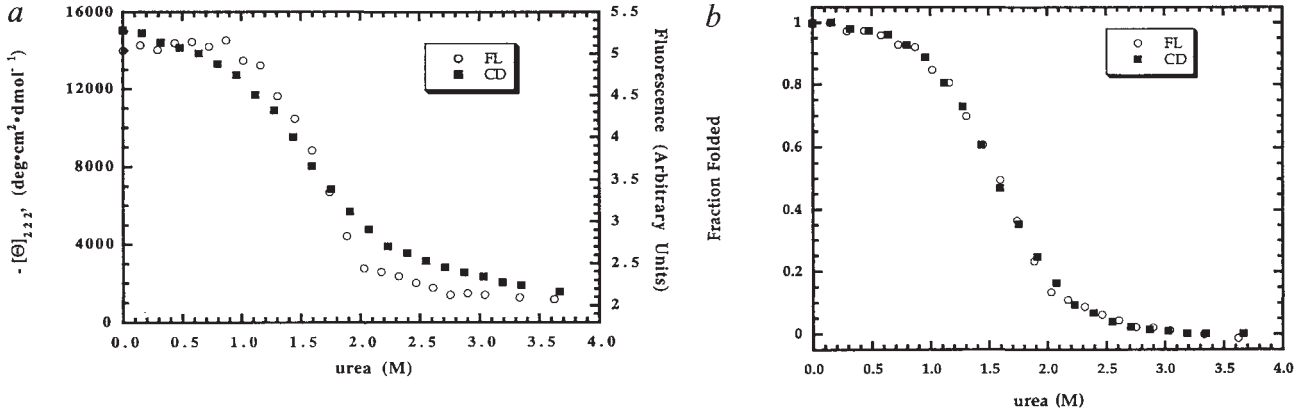

Fig. 2 CD and fluorescence data for urea-induced unfolding of wild-type apoMb intermediate (pH 4.2, 4 °C, 4 mM citrate): *a*, raw data and *b*, normalized unfolding curves.

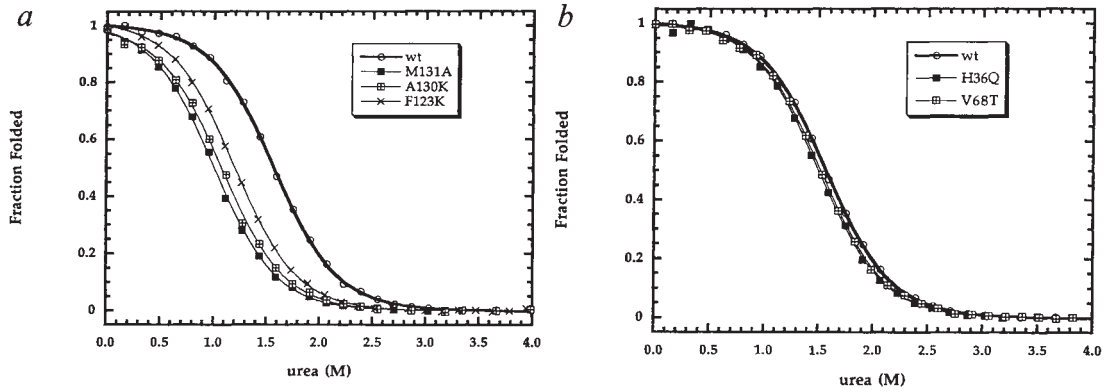

Fig. 3 Representative normalized urea unfolding curves measured by CD for the intermediate (pH 4.2, 4 °C, 4 mM citrate): *a*, M131A, A130K, F123K, and wild type and *b*, V68T, H36Q, and wild type.

apparent free energies of unfolding (ΔG_{IU}), when measured by fluorescence and CD, for wild type and each mutant studied (except A130L, see below). The average of the apparent unfolding free energies determined by CD and fluorescence is reported here as ΔG_{IU} (Table 1).

Analysis of native protein stability

The native state of apomyoglobin unfolds by urea denaturation in a broad, apparent two-state transition when measured by CD (Fig. 5)[20,21]. Examination of the urea unfolding transition by fluorescence, which is more sensitive to the presence of the intermediate, reveals, however, significant three-state unfolding character (Fig. 6). This increase in sensitivity stems from the intermediate's increased fluorescence relative to both N and U. To quantitatively measure the relative stability of N for apomyoglobin and its mutants, a three-state model was simultaneously fitted to CD and fluorescence data. The fluorescence data alone is not sufficient to accurately determine all the parameters of a three-state model, but the addition of CD data constrains the fit and reduces the error in the fitted parameters. To facilitate comparison of wild-type and mutant proteins, *m*-values (N→I and I→U) were fixed at wild-type values as described above. The sum of the free energies of the N→I and I→U transitions yields the stability of the native state (ΔG_{NU}) (Table 1).

Previous studies have used two-state urea-induced unfolding of the native state at pH 7.8 to measure native stability[1,21]. The fluorescence data shown here demonstrate that for wild type this assumption is nearly valid, but is much less so for destabilized mutants where the intermediate becomes significantly populated during unfolding (Fig. 6). Assuming two-state unfolding will result in an underestimate of the free energy of unfolding for destabilized mutants because the transition is broadened, leading to an apparent lowered *m*-value and exaggeration of mutant ΔGs for the N→U transition of mutants relative to wild type.

Effects of mutations on intermediate stability

Met 131 is the only residue in myoglobin to form large, conserved hydrophobic contacts with both the A and G helices. As a result, if native-like packing is important for the intermediate, this residue may be expected to act as an anchor for the AGH subdomain. The pH unfolding profile of M131A shows typical three-state behaviour (data not shown). The N→I pH midpoint is increased by 1.5 pH units over the wild-type protein, while the I→U transition is unchanged. In contrast, the urea denaturation curve of M131A at pH 4.2 shows a destabilization of the intermediate by 0.9 kcal mol^{-1} compared to wild type (Figs 2,3). The

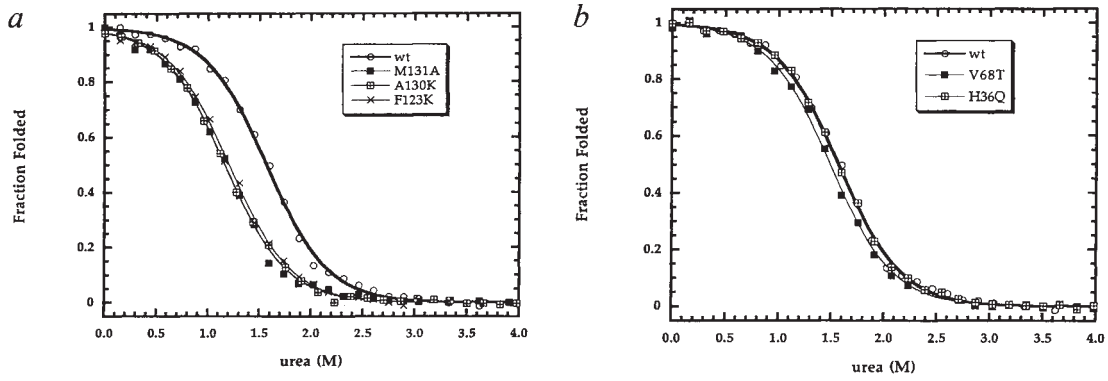

Fig. 4 Representative normalized urea unfolding curves measured by fluorescence for the intermediate (pH 4.2, 4 °C, 4 mM citrate): *a*, M131A, A130K, F123K, and wild type and *b*, V68T, H36Q, and wild type.

Table 1 Apparent stabilities of mutants (kcal mol^{-1})

Mutant	Location	Intermediate $-\Delta G_{IU}$[1] ($\Delta\Delta G_{wt}$)	Native $-\Delta G_{NU}$[2] ($\Delta\Delta G_{wt}$)
wild type		2.9	5.6
W7F	A/H interface	1.8 (1.1)	4.7 (0.9)
W14F	A/E interface	2.4 (0.5)	4.5 (1.1)
M131A	A/G/H interface	2.0 (0.9)	3.4 (2.2)
F123K	GH loop	2.2 (0.7)	3.5 (2.1)
A130K	A/H interface	2.0 (0.9)	3.5 (2.1)
A130L	A/H interface	3.0[3] (-0.1)	4.7 (0.9)
H36Q	BC/G interface	2.8 (0.1)	4.8 (0.8)
V68T	E helix	2.7 (0.2)	5.0 (0.6)

[1]Stability of the intermediate to urea-induced unfolding at pH 4.2, 4 °C, 4 mM citrate; the change from wild type is given in parentheses.
[2]Stability of native apoMb to urea-induced unfolding at pH 7.8, 4 °C, 10 mM Hepes; the change from wild type is given in parentheses.
[3]This mutant, in contrast to the other mutants in this table gives slightly different results when the transition curve is measured either by CD or by fluorescence (see Fig. 7). The average of the fluorescence (2.9 kcal mol^{-1}) and CD (3.1 kcal mol^{-1}) results is given here.

native state of M131A (at pH 7.8) is also severely destabilized ($\Delta\Delta G$=2.2 kcal mol^{-1}; Figs 5,6).

Ala 130 is nestled in the middle of the A/H helix interface. Hughson *et al.*[1] found that the native state of A130K was severely destabilized (ΔpH_{mid}=0.8), while I was nearly unaffected (ΔpH_{mid}=0.1). Analysis of this mutant by urea denaturation reveals significant destabilization of the intermediate (0.9 kcal mol^{-1}) in addition to severe destabilization of the native state (2.1 kcal mol^{-1}). The mutant A130L was found by Hughson *et al.* [1] to slightly stabilize I against acid denaturation (ΔpH_{mid}=0.1). Their result is confirmed here by urea denaturation using CD ($\Delta\Delta G_{IU}$=0.2 kcal mol^{-1}) but not when using fluorescence ($\Delta\Delta G_{IU}$=0) to monitor unfolding (Fig. 7). This small difference in stability measured by the two probes probably results from increased helicity in the unfolded form, which is apparent in the raw CD data (not shown) or the presence of an additional intermediate. The fact that the A130L mutation does not destabilize the intermediate indicates that it is loosely packed; probably the increased nonpolar surface area of the leucine side chain compensates for the packing defect. The significant destabilization of the native state by this mutant (0.9 kcal mol^{-1}) demonstrates that the hydrophobic core is tightly packed in the native state.

Phe 123 is located in the loop connecting the G and H helices and is extremely well conserved among myoglobins[22]. Its side chain protrudes into the GH hydrophobic interface, contacting residues in both the G and H helices. Hughson *et al.*[1] showed that F123K significantly destabilizes both N (0.8 pH units) and I (0.4 pH units). These results are confirmed here with F123K showing a destabilization of 2.1 and 0.7 kcal mol^{-1} for N and I respectively.

Trp 7 is buried deeply in the A/H helix interface. W7F is a relatively conservative mutation which occurs in many natural myoglobin sequences[22]. Despite this conservation, the native state is significantly destabilized by W7F ($\Delta\Delta G$=0.9 kcal mol^{-1}). In addition, the intermediate is significantly destabilized (1.1 kcal mol^{-1}). The analysis of fluorescence data for this mutant is somewhat complicated by the fact that one of the two Trp probes is being removed. The continued similarity of the normalized CD and fluorescence unfolding transitions supports, however, the assumption that both Trp residues report on the same unfolding event (that is, both Trps are probes of tertiary structure). Additional evidence for this assumption comes from W14F, a mutation in the A/E helix interface. This mutation also significantly destabilizes the native state (1.1 kcal mol^{-1}), but has only a moderate effect on the stability of I (0.5 kcal mol^{-1}). Again, the CD and fluorescence data for this mutant yield similar unfolding curves. The moderate loss of I stability caused by W14F is somewhat surprising given that the E helix is not formed in the intermediate[9,11]. Careful examination of the myoglobin crystal structure[23] reveals, however, that Trp 14 is located along the hydrophobic stripe of the amphipathic A helix, projecting towards the G and H helices. This location would be expected to contribute to nonspecific hydrophobic interactions stabilizing the intermediate.

Control mutations

In order to confirm that interactions outside the AGH subdomain are not important to the intermediate's stability, two mutations outside this subdomain were examined. His 36 in the B/C helix border interacts with Phe 106 in native myoglobin. His 36 also forms an electrostatic interaction with Glu 38[24]. Disruption of these interactions by H36Q results in significant destabilization of the native state (0.8 kcal mol^{-1}, ΔpH_{mid}=1.5[20]) while the intermediate's stability is unaffected. This result is consistent with the findings of Barrick and Baldwin[20] and indicates a lack of involvement of the B/C helix border in the intermediate. Mutation of Val 68, in the E helix, to Thr results in a moderate destabilization of the native state (0.6 kcal mol^{-1}), but little change in the intermediate's stability, confirming the lack of involvement of the E helix in the intermediate, consistent with the findings of Olson and co-workers[25].

Implications

In this study we have re-examined the role of specific and nonspecific hydrophobic packing in the stabilization of the intermediate. The use of urea denaturation and combined CD and fluorescence analysis allows an accurate quantitative measurement of the effects of mutation on both N and I, revealing important packing effects in I that were previously undetected. Mutations that reduce hydrophobicity of side chains (M131A, W7F, W14F) or introduce a charged group into the AGH hydrophobic interface (A130K, F123K) dramatically destabilize the intermediate. In contrast, the introduction of a larger hydrophobic side chain (A130L) has little or no effect on the stability of the intermediate. All packing mutants in the AGH subdomain studied here by urea denaturation (except A130L) show significant effects on the stability of the intermediate, indicating that specific residues make energetically important packing interactions. These results support a model in which combined nonspecific

hydrophobic interaction and partially formed native like interactions provide the stability of the folding intermediate. The discovery of specific packing in this intermediate helps to explain how folding intermediates can acquire native-like tertiary folds in their structured subdomains[26–28] and is consistent with molecular dynamics simulations[29] predicting specific tertiary contacts in I.

Comparison of the energetic effects of mutation on the native species and the intermediate reveals the extent of specific packing acquired in the intermediate. The ratio of $\Delta\Delta G_{NU}$ to $\Delta\Delta G_{IU}$ is ~2, for mutants targeting interactions largely or wholly within the AGH subdomain (Table 1) within the AGH subdomain (Table 1), implying approximately half strength interactions in the intermediate compared to the native state. This ratio is similar for most mutants involving packing wholly within the AGH subdomain, further supporting the view of weakened but native-like packing. The quantitatively different intermediate destabilizing effects of Trp 14 in the A/E helix interface ($\Delta\Delta G$=0.5 kcal mol^{-1}) and Trp 7 in the A/H helix interface ($\Delta\Delta G$=1.1 kcal mol^{-1}) give a rough indication of the relative importance of nonspecific hydrophobic (within the AGH subdomain) and specific packing interactions in stabilizing the intermediate. Marmorino and Pielak also find a stabilizing tertiary interaction in the iso-1-ferricytochrome *c* intermediate, but, in contrast to apoMb, its strength is the same in both the intermediate and native form[30], indicating a more tightly packed intermediate.

Constancy of intermediate structure

All of the mutant intermediates in this study have similar secondary structure content as measured by CD (–13,500–15,000 deg cm^2 dmol^{-1}) (data not shown). Although the stability of the intermediate can be easily diminished by mutation, its helical structure is relatively constant. The fluorescence intensities of mutant intermediates are much more variable, varying ±15% from wild type, but within these limits are also similar for all of the mutants studied (data not shown). As expected, maintenance of full secondary structure is also observed for the native state. Previously, Kallenbach and co-workers reported significant changes in the ellipticities of the native and intermediate conformations for different mutants[31]. In their study CD was measured under a single set of conditions, however, both for N (pH 7, 25 °C) and I (pH 3 + 0.5M NaCl, 25 °C), rather than measuring the entire unfolding curve. Since mutants affect stability, the N or I species may not be fully populated under these conditions for all mutants. Measuring the complete unfolding transitions of mutants ensures that the fully populated N or I species is observed, and this procedure gives a quantitative measurement of the change in stability.

Cooperativity of intermediate unfolding

The data presented supports a two-state unfolding mechanism for the intermediate, but cannot prove it. Stronger evidence of an approximate two-state mechanism comes from the kinetic studies of Jamin and Baldwin (unpublished results). Although the two-state nature of the unfolding transition cannot be proven with the current data, the high cooperativity of the unfolding is readily apparent. Mutants in the G/H helical interface influence the stability of the intermediate measured by Trp-monitored unfolding. Since both Trp residues are in the A helix, distant from the mutation site, the destabilization caused by G/H mutations ripples through the entire structure. This effect at a distance is the hallmark of a cooperatively folded structure. Additional evidence for cooperativity stems from the coincidence of probes of secondary (CD) and tertiary structure (fluorescence). Fluorescence measures the tertiary environment of Trp residues as indicated by large fluorescence changes during the transition from N to I and I to U. These changes reflect the different degrees of hydrophobic burial of the Trp probes and the distance of specific quenching groups from them. On mutation, the unfolding curves monitored by CD and fluorescence remain similar, demonstrating the strong link between secondary and tertiary structure (Fig. 7*b*). This cooperativity observed in the apomyoglobin intermediate contrasts with the non-cooperative behaviour seen in mutations of the α-lactalbumin molten globule, where proline mutations unfold only individual helices (B.A. Schulman and P.S. Kim, personal communication).

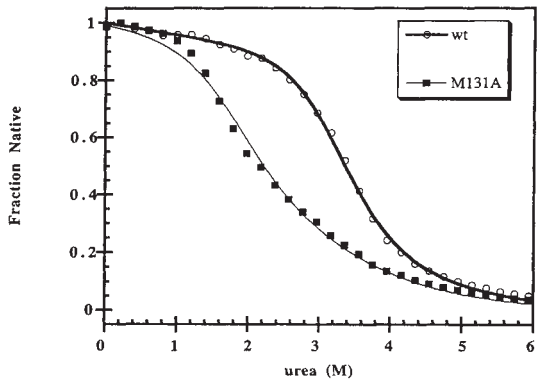

Fig. 5 Representative urea unfolding curves for native state (pH 7.8, 4 °C, 10 mM Hepes) measured by CD. Wild type and M131A.

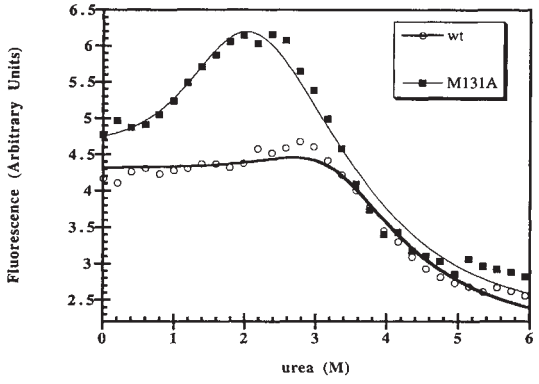

Fig. 6 Representative urea unfolding curves for native state (pH 7.8, 4 °C, 10 mM Hepes) measured by fluorescence. Wild type and M131A.

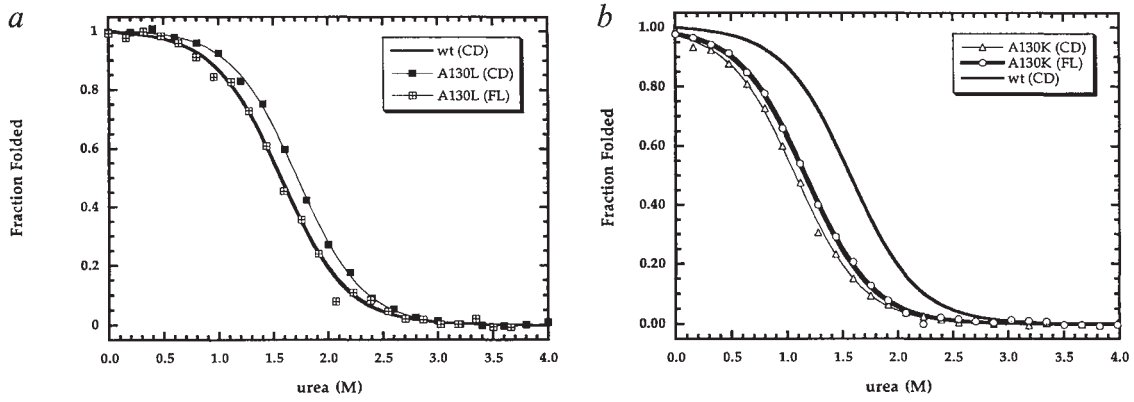

Fig. 7 Normalized CD and fluorescence urea unfolding curves for intermediate (pH 4.2, 4 °C, 4 mM citrate): *a*, A130L and *b*, A130K.

Why do acid and urea induced unfolding give different results?

Conflicting results observed here by urea denaturation and previously by pH denaturation are unexpected and point to important differences in the nature of acid and urea unfolding. One significant difference between urea and acid unfolding is the unfolded form itself. The acid-unfolded protein maintains some residual ellipticity (–3000 deg cm^2 $dmol^{-1}$), whereas the urea-denatured form has virtually none[8,16]. Goto and co-workers have shown that the acid-denatured form has a smaller radius of gyration than the urea-unfolded protein[10]. They also observe that the acid-unfolded state has no significant *m*-value associated with its urea-induced unfolding[8] and does not bind the hydrophobic dye ANS[6], reflecting a lack of organized structure in this state. This difference between the urea- and acid-unfolded states seems unlikely to play a major role in explaining the data, as all specific packing is likely to be lost in both the urea- and acid-unfolded states.

An alternative explanation is that pH denaturation is simply less sensitive to changes in the stability of I than urea unfolding, so that small changes in stability are subsumed by the error in pH measurements. In the three-state model of Barrick and Baldwin[16], the pK_as of the ionizing groups inducing denaturation of the intermediate are assumed to be infinitely low (that is, the groups are never protonated). According to this model, the pH midpoint of the I→U transition is sensitive to changes in the stability of the intermediate[2]. This prediction is in marked contrast to the experimental results observed in this study and the study of Hughson *et al.*[1]. This model has the convenience of introducing the minimum number of parameters to be determined. Because the intermediate has a more fluctuating structure than the native state it is unlikely, however, that any ionizable groups are prevented from being protonated in I. The native state has a more fixed structure, enabling it to generate large pK_a perturbations in sidechain pK_as[32]. Consequently, the assumption that ionizable groups cannot be protonated in one state, while perhaps valid for N→I, is less valid for I→U. When the model is made more realistic by including explicit pK_as both in I and U, the sensitivity of the pH_{mid} value decreases to changes in ΔG_{IU} (calculated curves not shown). More work is needed to determine if this explanation is correct for the striking difference between the results of pH and urea unfolding.

Further studies

The results presented here demonstrate the importance of specific hydrophobic packing in the folding intermediate of sperm whale apomyoglobin. This finding helps to explain much of the intermediate's stability and gives a qualitative measure of the relative strength of these tertiary interactions in the intermediate *versus* the native state: the changes in stability are about one-half as large in the intermediate. The method presented here allows sensitive measurement of the stability of the intermediate, which is critical for probing for other sources of stability, searching for non-native contacts in I, and understanding the nature of the I→U transition. It will be particularly interesting to use the methods described here to examine higher order folding intermediates[8,9] and hopefully to trace the evolution of native-like packing through stages in folding. The unexpected sensitivity of the intermediate to mutation means that solubilizing mutations may disrupt potential nascent structures in model peptide studies. For example, a polypeptide encompassing the G and H helices[33] contains the mutation M131K for solubilization, which would severely destabilize the intermediate and probably also any fragile structure formed by the GH polypeptide. The lack of detectable structure in this model may be caused either by the mutation or by insufficient stability of the fragment. Interestingly, F123K is the only mutant studied that has a significant perturbation of the I→U pH transition midpoint . This shift is probably indicative of a direct effect of this mutation on the pK_a of a residue involved in the I→U transition. Detection of similar mutants altering the I→U pH midpoint may help in mapping the groups responsible for the I→U transition.

Methods

Construction of mutants. Mutants were constructed from the synthetic myoglobin gene from pMb413b (originally a gift from S.G. Sligar[34], University of Illinois). W7F and W14F plasmids were gifts from Carol Rohl and all other mutations were made as described[1,20] or by recombinant polymerase chain reaction[35]. DNA sequences were

verified by standard dideoxy methods. Protein was expressed in *Escherichia coli* (TB1) as described[1]. All DNA manipulations were performed using standard techniques[36].

Protein purification. Proteins were purified essentially as described[9], but without the addition of exogenous haem. Apoprotein was prepared using the acid-acetone method[37]. Purity was >95% as judged by SDS-PAGE. Proteins were extensively dialysed against water after haem removal and lyophilized for storage. Protein stocks in water were stored at –20 °C and used within 36 hours of thawing. Apoprotein concentrations were measured by UV absorbance in 6 M guanidine HCl, as described by Edelhoch[38].

CD and fluorescence. CD data were collected on an Aviv 60DS circular dichroism spectropolarimeter at 4 °C. All measurements were taken in a 1 cm path length cuvette at 222 nm with 1–2 μM apoMb. Fluorescence measurements were taken on a SLM-Aminco Bowman Series 2 Luminescence Spectrometer in a 1 cm×0.5 cm cuvette with 1 μM apoMb thermostatted to 4 °C. Excitation was at 280 nm, with fluorescence emission measured at 320 nm. $[\theta]_{222}$ was constant up to 80 μM apoMb for both N and I and the stability of I did not vary with concentration up to 100 μM, as previously described[9], indicating that any aggregation does not affect these measurements of stability at the apoMb concentrations used in this study.

Analysis. All curve fitting was performed using KaleidaGraph (Abelbeck Software). Urea denaturation curves of I were fitted to a two-state equation with linear baselines as described above. Free energies of unfolding were calculated using the linear extrapolation model with baselines calculated using the entire data set[17]. Reported free energies of unfolding of I are the average of the CD and fluorescence measurements.

Urea denaturation of the native state was fitted to a three-state equation as described by Barrick and Baldwin[1] (their eq. 15). A pooled linear native baseline (average of all mutants studied) was applied to the CD measurements. No baselines were applied to the fluorescence data, as these baselines were close to zero and could not be determined accurately by the data. The fraction of native ellipticity for the CD data was fixed at 1 for N, 0.7 for I, and 0 for U. Fluorescence values for N, I and U were permitted to float during the fitting because they vary with mutation. The reported free energies of unfolding of N are the result of a simultaneous fit of the CD and fluorescence data to the three-state model.

Received 15 January; accepted 15 March 1996.

Acknowledgements
We thank S. Loh and C. Rohl for extensive advice, assistance, discussion and critical review of this work, particularly the discussion of the sensitivity of pH midpoint shifts to changes in stability. M.S.K. 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.

1. Hughson, F.M., Barrick, D. & Baldwin, R.L. Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry* **30**, 4113–4118 (1991).
2. Barrick, D. & Baldwin, R.L. The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Sci.* **2**, 869–876 (1993).
3. Griko, Y.V., Privalov, P.L., Venyaminov, S.Y. & Kutyshenko, V.P. Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol.* **202**, 127–138 (1988).
4. Cocco, M.J. & Lecomte, J.T. The native state of apomyoglobin described by proton NMR spectroscopy: interaction with the paramagnetic probe HyTEMPO and the fluorescent dye ANS. *Protein Sci.* **3**, 267–281 (1994).
5. Cocco, M.J. & Lecomte, J.T. Characterization of hydrophobic cores in apomyoglobin: a proton NMR spectroscopy study. *Biochemistry* **29**, 11067–11072 (1990).
6. Goto, Y. & Fink, A.L. Phase diagram for acidic conformational states of apomyoglobin. *J. Mol. Biol.* **214**, 803–805 (1990).
7. Griko, Y.V. & Privalov, P.L. Thermodynamic puzzle of apomyoglobin unfolding. *J. Mol. Biol.* **235**, 1318–1325 (1994).
8. Nishii, I., Kataoka, M. & Goto, Y. Thermodynamic stability of the molten globule states of apomyoglobin. *J. Mol. Biol.* **250**, 223–238 (1995).
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. Kataoka, M. *et al.* Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering. *J. Mol. Biol.* **249**, 215–228 (1995).
11. Hughson, F.M., Wright, P.E. & Baldwin, R.L. Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249**, 1544–1548 (1990).
12. 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).
13. Weaver, D. Hydrophobic interaction between globin helices. *Biopolymers* **32**, 477–490 (1992).
14. Richmond, T.J. & Richards, F.M. Packing of alpha-helices: geometrical constraints and contact areas. *J. Mol. Biol.* **119**, 537–555 (1978).
15. Waltho, J.P., Feher, V.A., Merutka, G., Dyson, H.J. & Wright, P.E. Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin. *Biochemistry* **32**, 6337–6347 (1993).
16. Barrick, D. & Baldwin, R.L. Three-state analysis of sperm whale apomyoglobin folding. *Biochemistry* **32**, 3790–3796 (1993).
17. Santoro, M.M. & Bolen, D.W. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988).
18. Pace, C.N., Shirley, B.A. & Thomson, J.A. in *Protein Structure* (ed. Creighton, T.E.) 311–330 (IRL Press, New York, 1992).
19. Ramsay, G., Ionescu, R. & Eftink, M.R. Modified spectrophotometer for multi-dimensional circular dichroism/fluorescence data acquisition in titration experiments: application to the pH and guanidine-HCl induced unfolding of apomyoglobin. *Biophys. J.* **69**, 701–707 (1995).
20. Barrick, D., Hughson, F.M. & Baldwin, R.L. Molecular mechanisms of acid denaturation. The role of histidine residues in the partial unfolding of apomyoglobin. *J. Mol. Biol.* **237**, 588–601 (1994).
21. Kiefhaber, T. & Baldwin, R.L. Intrinsic stability of individual alpha helices modulates structure and stability of the apomyoglobin molten globule form. *J. Mol. Biol.* **252**, 122–132 (1995).
22. Bashford, D., Chothia, C. & Lesk, A.M. Determinants of a protein fold. Unique features of the globin amino acid sequences. *J. Mol. Biol.* **196**, 199–216 (1987).
23. 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).
24. Garcia, M.B., Chen, L.X., March, K.L., Gurd, R.S. & Gurd, F.R. Electrostatic interactions in sperm whale myoglobin. Site specificity, roles in structural elements, and external electrostatic potential distributions. *J. Biol. Chem.* **260**, 14070–14082 (1985).
25. Hargrove, M.S. *et al.* Stability of myoglobin: a model for the folding of heme proteins. *Biochemistry* **33**, 11767–11775 (1994).
26. Wu, L.C., Peng, Z.Y. & Kim, P.S. Bipartite structure of the alpha-lactalbumin molten globule. *Nature Struct. Biol.* **2**, 281–286 (1995).
27. Peng, Z.Y., Wu, L.C., Schulman, B.A. & Kim, P.S. Does the molten globule have a native-like tertiary fold? *Phil. Trans. R. Soc. Lond.* B **348**, 43–47 (1995).
28. Morozova, L.A., Haynie, D.T., Arico-Meundel, C., Dael, H.V. & Dobson, C.M. Structural basis of the stability of a lysozyme molten globule. *Nature Struct. Biology* **2**, 871–875 (1995).
29. Tirado-Rives, J. & Jorgensen, W. Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry* **32**, 4175–4184 (1993).
30. Marmorino, J.L. & Pielak, G.J. A native tertiary interaction stabilizes the A state of cytochrome c. *Biochemistry* **34**, 3140–3143 (1995).
31. Lin, L., Pinker, R.J., Forde, K., Rose, G.D. & Kallenbach, N.R. Molten globular characteristics of the native state of apomyoglobin. *Nature Struct. Biol.* **1**, 447–452 (1994).
32. Cocco, M.J., Kao, Y.H., Phillips, A.T. & Lecomte, J.T. Structural comparison of apomyoglobin and metaquomyoglobin: pH titration of histidines by NMR spectroscopy. *Biochemistry* **31**, 6481–6491 (1992).
33. Shin, H.C. *et al.* Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. *Biochemistry* **32**, 6356–6364 (1993).
34. Springer, B.A. & Sligar, S.G. High-level expression of sperm whale myoglobin in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **84**, 8961–8965 (1987).
35. Landt, O., Grunert, H.P. & Hahn, U. A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* **96**, 125–128 (1990).
36. Sambrook, J., E.F. Fritsch & Maniatis., T. *Molecular cloning : a laboratory manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y, 1989).
37. 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).
38. Edelhoch, H. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**, 1948–1954 (1967).
39. Kraulis, P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).