

## Intrinsic Stability of Individual $\alpha$ Helices Modulates Structure and Stability of the Apomyoglobin Molten Globule Form

Thomas Kiefhaber and Robert L. Baldwin\*

Department of Biochemistry  
Stanford University School of  
Medicine, Stanford  
CA 94305, USA

During acid-induced unfolding of apomyoglobin, a partly folded form is observed at pH values of around four. In this form, the A, G and H helices are folded, while the rest of the molecule, including the B helix, demonstrates little structure. The partly folded form has been described as a molten globule form. To determine the factors that govern the structure and stability of this form, we introduced two helix-stabilizing mutations into the B helix, and tested their effect on the structure and stability of both the native form and the molten globule form. We show that the two Gly  $\rightarrow$  Ala replacements in the B helix produce altered fluorescence and CD properties of the partly folded intermediate, a result which implies that the B helix has become part of the structured region of the molten globule form. The helix content of a model peptide containing the sequence of the B helix is increased by the G  $\rightarrow$  A replacements, as is the helix content of the molten globule intermediate, whereas the stability and the helix content of the native protein are not altered. The observed increase in helicity is larger in the folding intermediate than in the model peptide, suggesting that nonspecific interactions, such as the hydrophobic interactions exhibited by the entire polypeptide chain, amplify the effect of intrinsic helix stability. The overall results suggest that the intrinsic stability of each individual helix is a factor in deciding whether or not that helix becomes part of the structured molten globule.

© 1995 Academic Press Limited

**Keywords:** molten globule; apomyoglobin; helix stability; B helix; G  $\rightarrow$  A replacements.

\*Corresponding author

### Introduction

Partially folded intermediates are sometimes observed in proteins under conditions in which tertiary contacts are weakened, e.g. after the removal of a ligand or at acid pH. These intermediates were named molten globules after their characteristic properties of a compact globular shape with a high degree of secondary structure and few, if any, specific tertiary interactions (for reviews see Kuwajima, 1989; Dobson, 1992; Ptitsyn, 1992). The molten globule form of apomyoglobin (apoMb) has been studied extensively. The acid-induced unfolding transition of apoMb is a three-state process at low salt concentrations (Griko *et al.*, 1988; Hughson & Baldwin, 1989; Barrick & Baldwin, 1993a). In a first

transition, a partially folded intermediate is formed at around pH 4. An acid-unfolded protein is produced at pH values below 3. The first transition, from the native (N) form to the molten globule form, is strongly influenced by the breakage of a hydrogen bond between His24 and His119, caused by protonation (Barrick *et al.*, 1994). The unfolding of the molten globule at low pH is thought to be caused by charge repulsion between the positive charges (Goto *et al.*, 1990). The intermediate (I) form shows about 35% helicity (compared to 55% in native apoMb), and an increased fluorescence emission of the Trp residues relative to both native and unfolded (U) forms. The structure of this intermediate has been investigated in detail, using amide hydrogen exchange in combination with 2D NMR analysis (Hughson *et al.*, 1990). The A, G and H helices are still present in the intermediate, as judged by the significant protection from amide proton exchange of the probes in these regions. Protons in helices C, D and E, and most protons in B, show little or no protection from exchange (protection factors close to 1), and are believed to be largely disordered in the

Present address: Thomas Kiefhaber, Biozentrum der Universität Basel, Abteilung Biophysikalische Chemie, Klingelbergstrasse 70, CH-4056 Basel, Switzerland.

Abbreviations: 2D, two-dimensional; apoMb, apomyoglobin; PCR, polymerase chain reaction; TCA, trichloroacetate; wt, wild-type; N, native; I, intermediate; U, unfolded.

molten globule. No information on the stability of the F helix is available because it has no slowly exchanging amide protons. Recent kinetic studies revealed that the A, G and H helices form very rapidly during folding at pH 6 (Jennings & Wright, 1993). They produce a transient intermediate that closely resembles the molten globule intermediate observed under equilibrium conditions. Later in the kinetic process of folding, a second kinetic intermediate is formed that has stable A, B, G and H helices. The remaining helices are formed in the final stage of folding, giving the native protein. In this work, we refer to the intermediate with stable A, G and H helices as  $I_1$ , and use  $I_2$  for the intermediate with stable A, B, G and H helices.

The factors that determine the stability and the structure of molten globule intermediates are still not well understood (see Lin *et al.*, 1994). We decided to test the possibility that the intrinsic stabilities of the individual helices are a key factor by introducing helix-stabilizing mutations into the B helix, which unfolds in the acid-induced N  $\rightarrow$  I transition of apoMb. The C, D and E helices do not show any protection from amide proton exchange in the molten globule (Hughson *et al.*, 1990), whereas some protons in the C-terminal region of the B helix are marginally protected. This may indicate that the B helix is on the verge of being stable in the molten globule intermediate; this theory is supported by the existence of a second kinetic folding intermediate, which has a stable B helix in addition to the A, G and H helices (Jennings & Wright, 1993). We therefore replaced amino acids at certain positions in the B helix, both in a model peptide and in the intact protein, in order to increase the helix stability of the isolated B helix, and to test the resulting effect on the structure and the stability of the molten globule form.

The molten globule intermediate of apoMb is known to be stabilized by binding certain anions, especially trichloroacetate ( $\text{TCA}^-$ ) (Goto *et al.*, 1990a). After this work was completed, the second kinetic intermediate  $I_2$  of apoMb was found to be stable at equilibrium in 20 mM NaTCA, as measured by the protection factors of protons in the various helices (Loh *et al.*, 1995). This finding indicates that the B helix, which is stable in  $I_2$  but not in  $I_1$ , is indeed on the verge of being stable in  $I_1$ . Also, Kataoka *et al.* (1995) have characterized, by solution X-ray scattering, the molten globule conformations of apoMb stabilized by binding  $\text{Cl}^-$  or  $\text{TCA}^-$ , respectively.

## Results

### Peptide models for the B helix

The B helix of apoMb is an amphipathic  $\alpha$ -helix. In native myoglobin, the hydrophobic side of the B helix faces the interior of the molecule and is in contact with the E helix, whereas the hydrophilic side is exposed to the solvent. According to the  $\alpha$ -helix propensity values determined in the model peptides, the two glycine residues at positions 23 and 25

wild-type: Ac-Y-D-V-A-G-H-G-Q-D-I-L-I-R-L-F-K-S-H-CONH<sub>2</sub>  
 G23A,G25A: Ac-Y-D-V-A-A-H-A-Q-D-I-L-I-R-L-F-K-S-H-CONH<sub>2</sub>

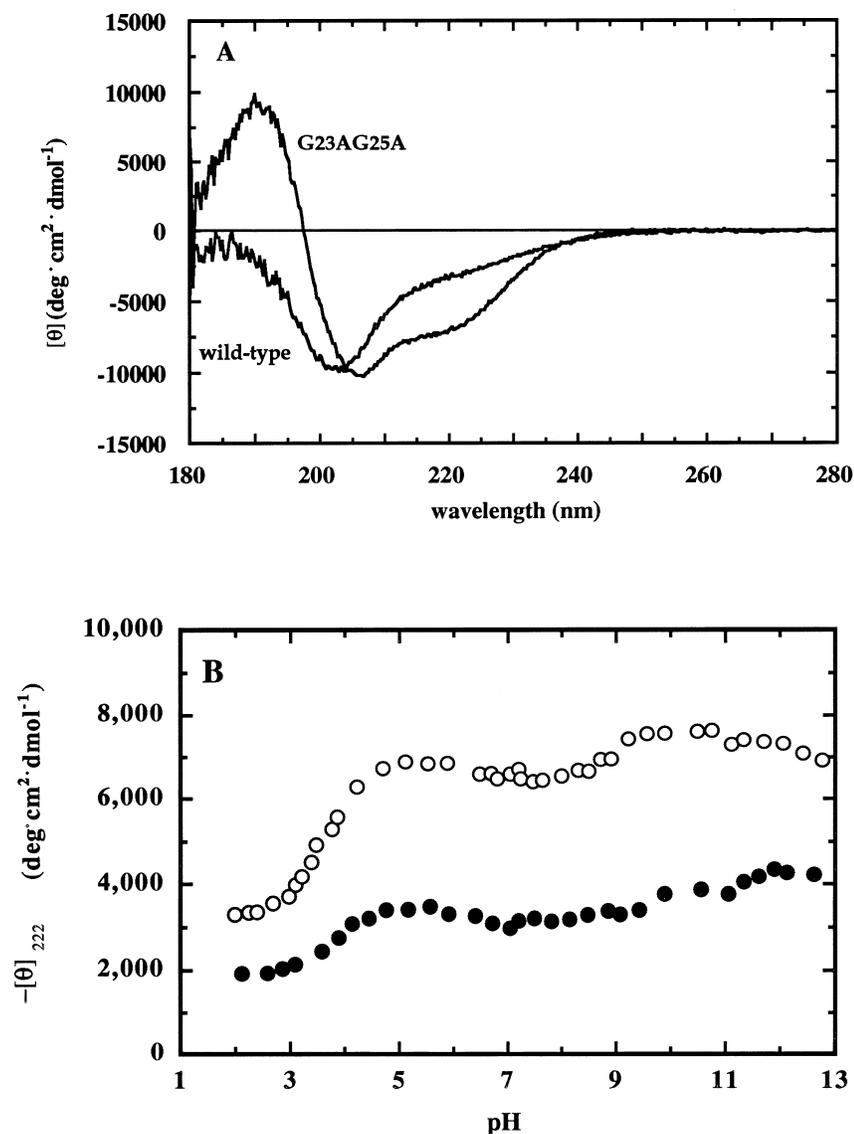
**Figure 1.** Sequences of the wild-type and G23AG25A peptides corresponding to the B helix of myoglobin.

should be highly unfavorable for the stability of the B helix (Chakrabarty *et al.*, 1991, 1994). We thus tried to stabilize the B helix by replacing the glycine residues with alanine (Figure 1), the best helix-forming amino acid in model peptides. Figure 2A shows CD spectra of the wild-type and G23AG25A peptides at neutral pH. The peptide containing the wild-type sequence exhibits a typical CD spectrum of a random coil conformation, whereas the peptide containing the two Gly  $\rightarrow$  Ala replacements shows a CD spectrum with distinctive minima at 222 and 208 nm. This is a clear indication of the presence of an  $\alpha$ -helical structure. Judged by the value of the ellipticity at 222 nm ( $\Theta = -6900 \text{ deg cm}^2 \text{ dmol}^{-1}$ ), the helicity of this peptide was approximately 20%.

The pH dependence of the peptide ellipticity shows that the helicity of the 2G  $\rightarrow$  A peptide is increased over the whole pH range, compared to the wild-type sequence (Figure 2B). Both peptides show a decrease in helicity between pH 3 and 4, with a midpoint at pH 3.6. This titration is probably caused by the protonation of one or both of the two Asp side-chains of the peptide. Minor signal changes corresponding to titration of the His and Arg side-chains are detectable around pH 7 and pH 11, respectively. These results indicate that the peptide corresponding to the wild-type B helix has some helical content, which is lost at pH values below 3. Comparison of the two peptides shows that the replacement of glycine residues 23 and 25 with alanine increases the helix content of the B helix peptide significantly.

### The stability of native apomyoglobin is nearly unchanged in the G23AG25A variant

Glycine 25 is a strongly conserved residue in all myoglobin and in nearly all hemoglobin sequences (position B6; see Stryer, 1988). It allows close contact between the B and the E helices (the E helix also has a glycine at E8 in the BE contact site). We therefore expected the disruption of tertiary interactions in the native form upon replacing this residue with the bulkier alanine residue. Introducing the G23A-G25A replacements into the polypeptide sequence of sperm whale myoglobin, however, yields native protein with the ability to bind heme. To compare the stability of the native form of wild-type apoprotein with that of the variant, we measured urea-induced unfolding curves. According to a three-state analysis of apoMb unfolding (Barrick & Baldwin, 1993a), the intermediate is only populated to a minor extent in the urea unfolding of native apoMb at pH 7.8. The unfolding curves monitored by fluorescence and by CD give almost identical results: fluorescence is more sensitive to the presence of the intermediate. Figure 3



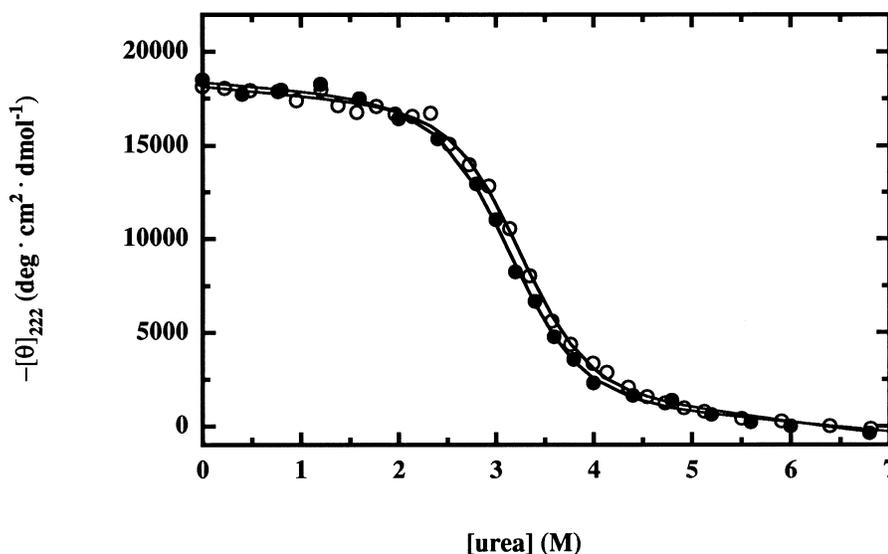
**Figure 2.** A, CD spectra of B helix peptides corresponding to the wild-type and G23AG25A sequences. Peptide concentration was 81  $\mu$ M in H<sub>2</sub>O. B, The pH-dependence of the negative ellipticity at 222 nm for peptides corresponding to wild-type ( $\bullet$ ) and G23AG25A B helix ( $\circ$ ). Peptide concentration was 17  $\mu$ M in 1 mM citrate/borate/phosphate (see Materials and Methods). All measurements were carried out at 0°C.

shows the urea-induced unfolding curves at pH 7.8 of the native form of the wild-type and the G23AG25A variant. In the absence of urea, both proteins show the same helical content, suggesting that the substitutions have not substantially changed the native structure. The two unfolding curves are similar: the urea molarity at the transition midpoint increases slightly, from 3.14 M for the wild-type to 3.25 M for the variant. The solid lines in Figure 3 were fitted by the two-state model. The apparent  $\Delta G^\circ$  values extrapolated to 0 M are  $-19.5$  kJ/mol for the wild-type and  $-21.6$  kJ/mol for the variant, and the respective values of  $m$  are 6.2 and 6.7 kJ mol $^{-1}$  M $^{-1}$ . Thus, the two proteins have similar stabilities in the native form.

These values should be contrasted with the results of a three-state fit, using more extensive data for pH-induced unfolding at various urea concentrations, and urea-induced unfolding at various pH values: for  $N \rightleftharpoons U$ ,  $\Delta G^\circ = -27.7$  kJ mol $^{-1}$  and  $m = 8.5$  kJ mol $^{-1}$  M $^{-1}$  (Barrick & Baldwin, 1993a, Table 1).

### The molten globule form of the G23AG25A variant has altered spectral properties

Figure 4 compares the acid-induced unfolding transitions of wild-type and G23AG25A apomyoglobin. The transitions were monitored by the decrease in negative ellipticity at 222 nm (Figure 4A), and by the change in fluorescence emission (Figure 4B), which is caused chiefly by the two Trp residues. apoMb also has three Tyr residues, but their fluorescence intensity at 320 nm should be very small. The wavelength of maximum emission is 336 nm for the wild-type and 335 nm for the mutant, compared to 347 nm of the acid-unfolded form. The pH midpoints of the  $N \rightarrow I$  and of the  $I \rightarrow U$  transitions are not affected by the amino acid replacements. Both the CD and the fluorescence measurements show altered spectral properties, however, of the intermediate formed by the G23AG25A protein. This spectral difference is more pronounced in the fluorescence-detected titration curve; 320 nm was chosen for monitoring the



**Figure 3.** Urea-induced unfolding curves of native wild-type (●) and G23AG25A (○) apomyoglobin in 10 mM Hepes (pH 7.8) at 0°C. Unfolding was monitored by the change in ellipticity at 222 nm in 1 cm cuvettes. The protein concentration was 1  $\mu$ M. The data were fitted by the two-state model (solid lines) with values of  $\Delta G(0) = -19.5$  kJ/mol and  $m = 6.2$  (kJ mol<sup>-1</sup> M<sup>-1</sup> for the wild-type protein and  $\Delta G(0) = -21.6$  kJ/mol with  $m = 6.7$  kJ mol<sup>-1</sup> M<sup>-1</sup> for the G23AG25A variant.

two transitions because the three forms, N, I and U, are most readily distinguished at this wavelength. The variant shows a strongly decreased yield in the Trp fluorescence emission at 320 nm. This indicates structural change in the intermediate near to at least one of the two Trp residues, caused by the replacement of glycines 23 and 25 by alanine. A slight increase in helical content of the molten globule intermediate is visible in the CD-detected pH titration. Both steps of the unfolding process lead to a decrease in the intensity of the CD signal, and the transitions from N  $\rightarrow$  I and from I  $\rightarrow$  U are not well separated. This makes it difficult to determine the CD properties of the molten globule intermediate under these conditions.

**Table 1.** Helix Stabilities of Peptides Corresponding to Myoglobin Helices

Peptide	$v^2w^{n-2a}$	$-[\Theta]_{222}^b$ (deg cm <sup>2</sup> dmol <sup>-1</sup> )
A <sub>wt</sub>	$3.7 \times 10^{-8}$	not helical <sup>c</sup>
B <sub>wt</sub>	$0.084 \times 10^{-8}$	3400 <sup>d</sup>
B (2G $\rightarrow$ A)	$87 \times 10^{-8}$	6900 <sup>d</sup>
G <sub>wt</sub>	$1.4 \times 10^{-8}$	<5800 <sup>e</sup>
H <sub>wt</sub>	$17 \times 10^{-8}$	9600 <sup>e</sup>

<sup>a</sup> The predicted equilibrium constant of the completely helical conformation calculated from helix propensities only, using data from Chakrabarty *et al.* (1994) and the Lifson-Roig equation:  $v$  is the nucleation parameter, and  $w^{n-2}$  is the product of the  $n-2$  helix propensities ( $w$ ) at internal positions in the helix.

<sup>b</sup> Measurements from peptides of the negative mean residue ellipticity at 222 nm. The value for 100% helix is approximately 34,000.

<sup>c</sup> Barrick & Baldwin, 1993b; pH 4.4.

<sup>d</sup> This paper (Figure 2B); pH 5.0.

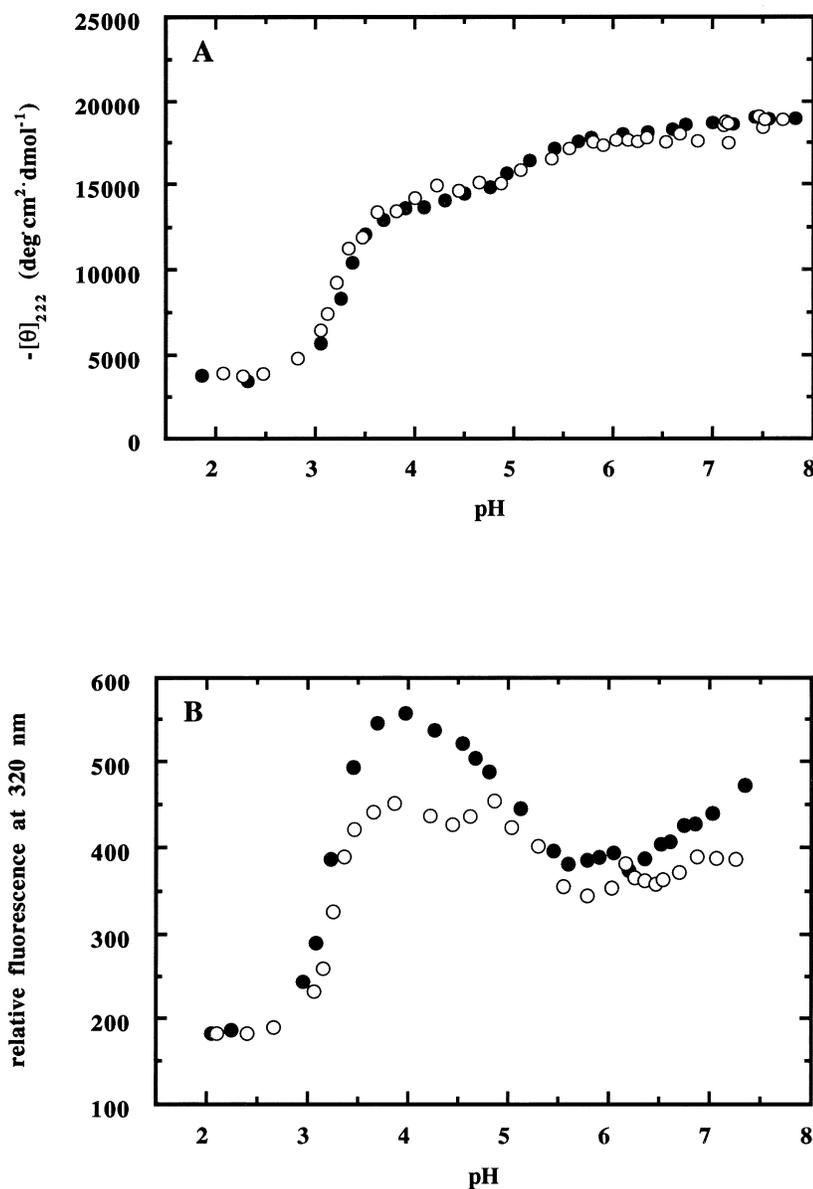
<sup>e</sup> Hughson *et al.*, 1991; pH 6.0. The value given for G<sub>wt</sub> is an upper bound, based on peptides containing some substitutions (see text).

### Stabilizing the B helix increases the helicity of the molten globule state

Molten globule intermediates of many proteins are stabilized in the presence of high concentrations of neutral salts, and in these conditions the acid-induced I  $\rightarrow$  U transition of apoMb is not observed (Goto *et al.*, 1990). In 0.5 M KCl, the acid-induced N  $\rightarrow$  I transition of apoMb occurs between pH 7 and pH 5 (Figure 5), and the molten globule intermediate is fully populated at pH values below 5. Figure 5 shows that in 0.5 M KCl, the pH midpoint of the N  $\rightarrow$  I transition is nearly identical for the wild-type and the variant, and the molten globule intermediates of both proteins are stable over a wide pH range. The negative ellipticity ( $\Theta$ ) of the variant intermediate is increased by about 3000 deg cm<sup>2</sup> dmol<sup>-1</sup> at 222 nm. These results show that both the helical content and fluorescence properties of the molten globule intermediate are significantly changed by stabilizing the B helix.

### Stability of the molten globule form towards urea-induced unfolding

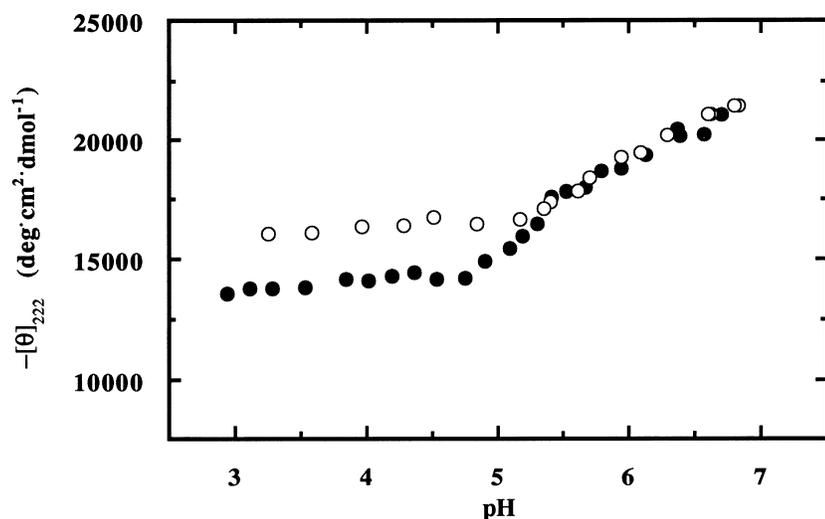
As shown above, the molten globule intermediate can be fully populated at low pH values in 0.5 M KCl. This enabled us to measure the stability of the intermediate towards unfolding by denaturant. Figure 6 shows urea-induced unfolding transitions of the wild-type and G23AG25A intermediates at pH 4.0 in 0.5 M KCl. The transitions were monitored by the CD signal at 222 nm. We used urea rather than GdmCl to induce unfolding, in order to avoid unwanted side-effects of the high chloride concentration introduced by using GdmCl. The curves show that the molten globule of the G23AG25A variant has an increased helicity compared to the wild-type



**Figure 4.** Comparison of the pH-induced unfolding of apomyoglobin wild-type (●) and the G23AG25A variant (○) in 2 mM citrate/citric acid at 0°C. A, Molar ellipticity at 222 nm. B, Transition monitored by the intensity of the tryptophan fluorescence emission at 320 nm after excitation at 278 nm. The protein concentration was 1 μM in all cases.

protein under all conditions. The slope of the unfolding transition of the G23AG25A molten globule is much shallower, however:  $m = 2.9 \text{ kJ mol}^{-1} \text{ M}^{-1}$ , compared to  $m = 5.5 \text{ kJ mol}^{-1} \text{ M}^{-1}$  for the wild-type intermediate, indicating less cooperative unfolding. Both transitions were fitted by the two-state model. The midpoint of the transition of the variant is increased by about 0.8 M. Linear extrapolation of the data to 0 M urea gives a decrease in stability of the molten globule form from  $-14.8 \text{ kJ/mol}$  for the wild-type to  $-10.3 \text{ kJ/mol}$  for the variant. This decrease is caused by the drastically reduced apparent  $m$ -value for the G23AG25A variant. In both transitions, the folded molten globule shows a sloping baseline, which is fitted automatically by the procedure of Santoro & Bolen (1988). The sloping baseline is more pronounced for the variant protein, and could reflect the pre-melting of some helical structure before the cooperative unfolding transition occurs. These findings suggest

that the two-state model is not valid for the urea-induced unfolding transition of the variant molten globule (see below). An alternative method of comparing the stabilities of the molten globule forms is to compare their  $\Delta G$  values in the transition region, where the equilibrium constants for both proteins can be measured under identical conditions. This method allows the comparison of the stabilities under identical solvent conditions: extrapolation to zero denaturant is not required. We assumed that 100% I is present at 0 M urea, and 0% I at 5.5 M urea. Comparison of the equilibrium constants for the wild-type and variant molten globules gives  $K_{\text{eq}}$  equal to 1.4 for the G23AG25A molten globule at the transition midpoint ( $K_{\text{eq}} = 1$ ) of the wild-type protein. This value corresponds to an increase in stability ( $\Delta\Delta G$ ) of  $-0.76 \text{ kJ/mol}$  for the variant molten globule. The midpoint of the unfolding transition is 0.4 M higher for the G23AG25A variant.



**Figure 5.** pH-induced unfolding of apomyoglobin wild-type (●) and the G23AG23A variant (○) in 2 mM citrate/citric acid at 0°C in 0.5 M KCl. The transition was monitored by the ellipticity at 222 nm. The protein concentration was 1 μM.

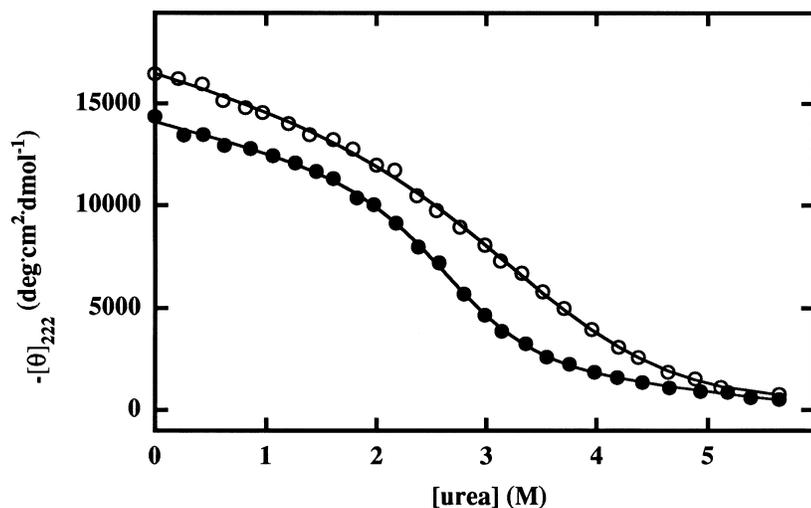
## Discussion

### Effects of glycine → alanine replacements on the helicity of the isolated B helix

The substitutions Gly23 → Ala and Gly25 → Ala have the following effects on the structure and stability of a short peptide containing the sequence of the B helix. Far-UV CD spectra, which measure helical content, indicate that the variant peptide adopts a partly  $\alpha$ -helical structure, whereas the wild-type B peptide shows a nearly random coil conformation. According to the intensity of the CD signal at 222 nm, the helical content is increased from about 10% to 20% at neutral pH by the G → A replacements. This increased helical content can be explained by the higher helix propensity of alanine compared to glycine (see below and Chakrabarty *et al.*, 1991, 1994). Both peptides show pH-dependent changes in the helical content, especially between pH 4 and pH 3, where the helical content drops drastically. This drop may be caused by the

protonation of aspartate 20 and/or 27. Asp20 is the N-terminal amino acid of the B helix peptide, and a negative charge in this position probably helps to stabilize the helix by interacting with the partial positive charge at the N terminus from the helix dipole. Protonation of Asp20 leads to a loss of this helix-stabilizing interaction. Interestingly, this strong decrease in the helix content of the peptide occurs in the same pH region where the apoMb molten globule unfolds. This might contribute to the loss of its stability between pH 4 and 3.

Table 1 gives the equilibrium constant for the completely helical conformation, as calculated from the Lifson-Roig theory and the helix propensities determined by Chakrabarty *et al.* (1994) for peptides with sequences corresponding to the wild-type A, B, G and H helices, and for the G23AG25A variant of the B helix. The calculated equilibrium constants are all extremely small ( $\sim 10^{-8}$ ), because these peptides have low helix contents, and the proportion of complete helix formed by each peptide is extremely small. Nevertheless, these equilibrium constants provide a useful measure of the relative intrinsic



**Figure 6.** Urea-induced unfolding of the molten globule of apomyoglobin wild-type (●) and the G23AG25A (○) variant in 0.5 M KCl. The conditions were; 10 mM citrate/citric acid (pH 4.0), 0.5 M KCl at 0°C. The unfolding was monitored by ellipticity at 222 nm. The data were fitted by the two-state model (solid lines), with  $\Delta G(0) = -14.8$  kJ/mol and  $m = 5.5$  kJ mol<sup>-1</sup> M<sup>-1</sup> for the wild-type and  $\Delta G(0) = -10.3$  kJ/mol and  $m = 2.9$  kJ mol<sup>-1</sup> M<sup>-1</sup> for the G23AG25A variant.

stabilities of the individual helices. The important result in Table 1 is that  $B_{wt}$  has by far the lowest intrinsic stability of the wild-type sequences for the A, B, G and H helices, but that the G23AG25A variant has a much higher intrinsic stability than  $B_{wt}$ , and has, in fact, the highest stability of this set of peptides.

The Lifson-Roig equation and experimental helix propensities can also be used to predict the actual helix contents of these peptides, but these numbers are not given in Table 1 for two reasons. First, the predicted values of fraction helix are in the range near 0.01, where they are not particularly sensitive to the values of the helix propensities, because the fraction helix is strongly influenced instead by the large number of partly helical conformations that are available. Second, the actual values of fraction helix for these peptides are larger than the predicted values by amounts that are fairly small in absolute value, but well outside experimental error. The nature of this discrepancy has not been investigated, but it is probably caused by nonpolar interactions and van der Waals interactions, such as the effects studied by Creamer & Rose (1995) via Monte Carlo simulations: see also Muñoz & Serrano (1994).

In Table 1, the helix contents (expressed as  $-\Theta_{222}$ ) are given for peptides with the wild-type sequences of helices B and H, but this cannot be done for helices A and G, whose wild-type sequences aggregate, when present as individual peptides. An upper bound was placed on the helix content of the G wild-type sequence by making a few substitutions that stop its aggregation, and increase its helix-forming tendency, as calculated from its helix propensities (Hughson *et al.*, 1991; see also Waltho *et al.*, 1993). The same procedure was used to investigate the helix-forming propensities of the isolated A helix sequence (Barrick & Baldwin, 1993b).

#### Effect of the glycine to alanine replacements on native apoMb

Introducing the G23A and G25A mutations into the protein does not influence the structure of native apoMb as judged by its far-UV CD and fluorescence properties. Also, the stability of the native form is not greatly affected. This almost unchanged stability is probably caused by the compensating effects of increased stability of the B helix, and decreased packing efficiency at the contact site between the B and E helices. Position B6 (glycine 25 in sperm whale myoglobin) is conserved in nearly all myoglobin and hemoglobin sequences, and allows close packing of the B and E helices. The fact that the stability of the native form of G23AG25A apoMb is almost unchanged allows us to assign changes in the pH-induced  $N \leftrightarrow I$  transition solely to structural changes in the intermediate itself. The helix content of native apoMb is not changed by the Gly  $\rightarrow$  Ala replacements. This was expected, since the B helix appears to be fully folded in native apoMb (see the protection factors measured by Hughson *et al.*, 1990),

and increasing its stability does not increase its helix content.

#### Effect of increased helix stability on the structure of the molten globule intermediate

In contrast to native apoMb, the structure of the molten globule form is affected by the increased stability of the B helix. The pH-induced  $N \rightarrow I$  transition of the G23AG25A protein shows altered far-UV CD and fluorescence properties between pH 4 and pH 5, where the molten globule intermediate is populated. This effect is pronounced in fluorescence emission of the Trp residues. The partly folded intermediate shows an increased fluorescence intensity compared to both native and unfolded apoMb, which makes it easy to detect the intermediate in unfolding transitions monitored by Trp fluorescence. This increase in fluorescence intensity is less pronounced in the G23AG25A variant than in the wild-type protein. The two Trp residues of myoglobin are both located in the A helix, and they probably monitor the interaction of the A, G and H helices in the molten globule. Significant changes in the fluorescence properties of the molten globule form should be indicative of structural changes in the ordered region of the molten globule.

The changes in the  $N \rightarrow I$  transition observed by CD at 222 nm are less pronounced, compared to the fluorescence measurements. A slight increase in helix content is visible around pH 4. Both the  $N \rightarrow I$  and the  $I \rightarrow U$  transitions show a decrease in the CD signal at 222 nm, and the transitions are not well separated. This makes it difficult to determine the CD properties of the molten globule form from the acid-induced unfolding transition curve.

As the KCl concentration is increased, the pH midpoint of the  $I \leftrightarrow U$  transition shifts to lower pH values (Goto & Fink, 1990), and in 0.5 M KCl, this transition is not observed: the molten globule intermediate is populated over a wide pH range. Both the helicity of the native form, and the midpoint of the acid  $N \rightarrow I$  transition, are identical in 0.5 M KCl for wild-type and G23AG25A apoMb, but the helicity of the molten globule intermediate is significantly increased in the variant. If the Gly  $\rightarrow$  Ala replacements only affect the helix content of the B helix in the intermediate, then the results indicate a 65% increase in the helicity of the B helix. Hydrogen exchange experiments on both the equilibrium molten globule (Hughson *et al.*, 1990) and on the first transient folding intermediate ( $I_1$ ) of wild-type apomyoglobin (Jennings & Wright, 1993) indicate some structure in the C-terminal region of the B helix. Thus, the Gly  $\rightarrow$  Ala replacements in the N-terminal part of the B helix may stabilize some existing helical structure in the region of the B helix. The helicity of the B helix peptide of the G23AG25A variant shows only partial helix formation, like the G helix and H helix peptides, and both the G and H helices are fully stable in the molten globule of the wild-type protein. This might suggest a direct

correlation between the intrinsic stability of the isolated helix and its ability to become fully stabilized in the molten globule.

Two other results suggest, however, that such a direct correlation is too simple. First, A helix peptide studies indicate that the isolated A helix peptide does not have observable helicity (Barrick & Baldwin, 1993b; see Table 1), whereas the A helix is fully stable in the molten globule. Second, the urea-induced unfolding curve of the G23AG25A molten globule is much less cooperative than that of the wild-type (Figure 6). These observations suggest that there is a qualitative difference between the mechanisms for stabilizing the B helix and the A, G and H helices in the variant molten globule.

### Cooperativity of folding in the molten globule

Whether the molten globule of apoMb is formed in a highly cooperative (two-state) reaction is not known for certain. A calorimetric study indicates that unfolding is not a first-order transition (Griko & Privalov, 1994). We found that urea-induced unfolding in 0.5 M KCl of the G23AG25A molten globule is not a two-state process. This was demonstrated by fitting the unfolding transition with the two-state model and obtaining two contradictory results when the stabilities of the wild-type and the G23AG25A variant are compared. The variant is found to be slightly more stable than the wild-type at the midpoint of the wild-type transition, but the variant is much less stable than the wild-type when their stabilities are extrapolated back to 0 M urea. The data suggest that the B helix of the G23AG25A variant begins to unfold at low urea concentrations, and its unfolding may be only weakly coupled to that of the A, G and H helices.

A possible approach for testing the cooperativity of the unfolding of the helical molten globule intermediates was introduced by Scholtz *et al.* (1995). They used the linear extrapolation method to analyze data for the urea-induced unfolding of a series of peptide helices of varying chain lengths. When  $s(M) = s(0) - mM/RT$ , where  $s$  is the average value of the helix propagation parameter of the Zimm-Bragg theory,  $m$  (for a single helical residue) is found to be  $96 \text{ J mol}^{-1} \text{ M}^{-1}$ . They proposed that, for the unfolding of a molten globule intermediate that obeys the two-state model,  $m = nm$ , where  $n$  is the number of helical residues in the molten globule species. For the apoMb intermediate  $I_1$  (with A, G and H helices),  $n = 61$  and  $nm = 5.9 \text{ kJ mol}^{-1} \text{ M}^{-1}$  versus  $5.5 \text{ kJ mol}^{-1} \text{ M}^{-1}$  observed here; for  $I_2$  (with A, B, G and H helices),  $n = 70$  and  $nm = 6.7 \text{ kJ mol}^{-1} \text{ M}^{-1}$  versus  $2.9 \text{ kJ mol}^{-1} \text{ M}^{-1}$  observed here. This comparison indicates that  $I_1$  does unfold by a two-state process, but the unfolding of  $I_2$  is less cooperative.

Ptitsyn & Uversky (1994) give a different method for investigating the cooperativity of unfolding. They

followed the procedure of Aune & Tanford (1969) by treating the interaction of a protein with a denaturant by the site binding model in which the sites are independent and equivalent. Then, provided that unfolding follows the two-state model, the slope of the unfolding curve at the denaturation midpoint can be simply related to  $\Delta v(\text{eff})$ , the difference between the number of denaturant molecules bound to the denatured form and to the native form. They assumed further that  $\Delta v(\text{eff})$  is a linear function of the protein molecular mass ( $M$ , in kDa) for all single domain proteins. They gave a correlation line  $\Delta v(\text{eff}) = 0.26 M + 1.05$  for data in the literature on the denaturant-induced unfolding of molten globule intermediates.

Our experimental values for the urea-induced unfolding of the apoMb molten globule form are  $\Delta v(\text{eff}) = 6.58$  (6.46) for the wild-type and 4.74 (4.93) for the G23AG25A variant. The first value is found by equation (2), and the value in parenthesis by equation (3) of Ptitsyn & Uversky (1994). If  $M$  in their correlation equation is taken to be the molecular mass of the whole protein (i.e. 17.3 for the expressed protein), then the predicted value for apoMb is  $\Delta v(\text{eff}) = 5.68$ , intermediate between the observed values for the wild-type and the variant. Only part of the protein is known to be structured in the folding intermediate, however, and the value of  $M$  in their correlation equation should correspond to the size of the structured region. Using values of  $M = 7.1$  and 8.1 for the structured regions in the wild-type (A, G and H helices) and in the G23AG25A variant (A, B, G and H helices), respectively, the predicted values of  $\Delta v(\text{eff})$  are 2.67 (wild-type) and 3.03 (variant). These predicted values are considerably smaller than the observed values, 6.5 (wild-type) and 4.8 (variant).

One possible explanation is that the results in the literature used by Ptitsyn & Uversky (1994) to fix their correlation line actually consist of a mixture taken from two-state and less cooperative unfolding reactions. As noted above, wild-type apoMb shows a two-state unfolding reaction by the tests applied here, but the G23AG25A variant shows a less cooperative unfolding reaction. It is also possible that one or more of the assumptions made by Ptitsyn & Uversky (1994) do not hold. For example, the sites for interaction with urea may not be uniformly distributed among all single-domain proteins according to their sizes. Use of the site-binding model for the interaction of urea and GdmCl with proteins has been criticized by Schellman (1978, 1990), on the grounds that these are weak interactions and a contact interaction, or solvent exchange, model would be more appropriate. Experimentally, the linear extrapolation method (Greene & Pace, 1974; Santoro & Bolen, 1988) is found to work better than the site-binding equation when  $\Delta G^\circ$  for unfolding is plotted against denaturant concentration, and the linear extrapolation method can be derived from the theory of weak interactions (Schellman, 1978).

## Implications for the mechanism of protein folding

We show here that, in the case of the B helix of apoMb, the stability of the isolated helix must reach a critical threshold value before the helix is stabilized by interacting with the ordered structure in the molten globule. A basic question concerning the molten globule intermediate of apoMb is: why are the A, G and H helices stabilized and the other Mb helices not stabilized in this intermediate? Part of the answer probably lies in the intrinsic stabilities of the isolated helices. It is important to note that nonspecific hydrophobic interactions between helices are expected to stabilize amphiphilic helices, and therefore helices that are found in native proteins (Baldwin, 1989). Evidence that helices can be stabilized by nonspecific hydrophobic interactions is provided by the observation that SDS stabilizes peptide helices (Nakano & Yang, 1981).

Experiments on *de novo* protein design provide a different view of how helices are stabilized in molten globule intermediates (Handel & De Grado, 1990; Raleigh & De Grado, 1992; Kamtekar *et al.*, 1993). Their results show that fixing the periodicity of hydrophobic and hydrophilic residues to allow the formation of amphiphilic helices is sufficient, in some cases, to produce four-helix bundle proteins, when residues specifying reverse turns are inserted at appropriate positions. These designed four-helix bundles can be highly stable, even though their stability results only from the burial of nonpolar surface, and the close packing characteristic of native proteins is not present. The results indicate that a molten globule species can have the tertiary fold of the corresponding native protein, even though the molten globule does not have fixed and specific tertiary interactions. Peng & Kim (1994) studied a molten globule species formed by the isolated helical domain of human  $\alpha$ -lactalbumin. They found that the correct disulfide bonds of native  $\alpha$ -lactalbumin are formed in the molten globule but not in the unfolded protein, a result which implies that the native tertiary fold is present in the molten globule species.

This second view of molten globule formation, and the one presented here, are complementary to each other. The experiments on *de novo* design achieve helix stabilization by ensuring that amphiphilic helices are formed that can interact loosely with each other and bury the nonpolar surface. Our experiments show that an increase in the stability of the B helix can be produced by increasing the helix propensities of individual residues through Gly  $\rightarrow$  Ala substitutions.

## Materials & Methods

### Peptide synthesis

Peptides were synthesized using F-moc chemistry on a Milligen 9050 synthesizer. The amino termini were acetylated with acetic anhydride. Peptide purification was

performed by C-18 reverse-phase chromatography (Pharmacia), and peptide identity and purity were measured by fast atom bombardment (FAB) mass spectroscopy. Observed molecular masses were within 0.2 mass units of the predicted values, and purity was greater than 99%.

### Mutagenesis

The plasmid pMb413b, containing a synthetic sperm whale myoglobin gene, was a gift from B. A. Springer and S. G. Sligar. Recombinant polymerase chain reaction (PCR) technology, according to Landt *et al.* (1990), was used for site-directed mutagenesis. In the first PCR reaction, a universal C-terminal primer and a mutagenic primer coding for the replacement of glycine residues at positions 23 and 25 by alanine were used to synthesize an oligonucleotide. This oligonucleotide was used in a second PCR reaction, together with a universal N-terminal primer. The two universal primers were homologous to regions just outside either end of the coding sequence, to give a final PCR product containing the complete sequence of the myoglobin gene. The product was cleaved with specific restriction enzymes, gel purified, and ligated into pMb413, from which the relevant region had been removed in a previous step.

### ApoMb preparation

Recombinant apoMb was expressed as described by Hughson *et al.* (1991), and purified according to Barrick and Baldwin (1993a). Heme was removed after purification by 2-butanone extraction (Teale, 1959; Hughson *et al.*, 1991). The concentration of apoMb was determined in 6 M GdmCl (pH 6.5), 20 mM sodium phosphate, according to the method of Edelhoch (1967), using molar extinction coefficients of  $\epsilon_{280} = 15,200$  and  $\epsilon_{288} = 10,800$ .

### CD measurements

CD data were collected on an Aviv 60DS spectropolarimeter, using cuvettes with a path length of 1 cm. The concentration of apoMb used was 1  $\mu$ M in 2 mM citrate buffer. CD measurements were carried out in 1 mM citrate/borate/phosphate buffer, with peptide concentrations of 17  $\mu$ M for the pH titration curves and 90  $\mu$ M for the spectra. The helicity of the peptides was independent of peptide concentration. All CD measurements were carried out at 0°C.

### Fluorescence studies

Fluorescence measurements were performed on a Perkin Elmer Lambda 5 spectrofluorimeter, using fluorescence cells of 1  $\times$  1 cm pathlength. Excitation was at 278 nm (3 nm spectral width), and fluorescence emission was monitored at 320 nm (5 nm spectral width). The protein concentration used was 1  $\mu$ M, at a temperature of 0°C.

### Stability towards unfolding with urea

Urea-induced unfolding curves were analyzed by fitting them as two-state transitions. The data were analyzed by the following linear extrapolation model:  $\Delta G^\circ(M) = \Delta G^\circ(0) - mM$ , where  $\Delta G^\circ$  is the standard change in Gibbs energy of unfolding, M is the urea molarity, and the slope  $m$  is proportional to the increase, produced

by protein unfolding, of the number of sites on the protein that interact with urea (Schellman, 1978). Thus,  $m$  is approximately proportional to the increase in the solvent-accessible surface area upon unfolding. The complete unfolding curves were fitted according to the equation given by Santoro and Bolen (1988), which sets the baselines by using data inside the transition zone as well as outside it. The unfolding transitions of the molten globule intermediates of wild-type and G23AG25A apoMb were also analyzed by comparing their stabilities in the unfolding transition according to:

$$\Delta\Delta G(M') = -RT \ln(1/K(M')) \quad (1)$$

where  $K(M')$  is  $[I]/[U]$  for the variant protein at the midpoint ( $M'$ ) of the unfolding transition of the wild-type protein ( $U$  is the unfolded protein and  $I$  is the partly folded intermediate), and  $\Delta\Delta G(M')$  is the difference in  $\Delta G^\circ$  between the variant and wild-type proteins.

## Acknowledgements

We thank Carol Rohl for advice and discussion and Antoine Firmenich and Helge Zieler for help with the site-directed mutagenesis. This work was supported by NIH Grant GM 19988. T.K. was a fellow of the Deutsche Forschungsgemeinschaft.

## References

- Aune, K. C. & Tanford, C. (1969). Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II. Dependence on denaturant concentration at 25°. *Biochemistry*, **8**, 4586–4590.
- Baldwin, R. L. (1989). How does protein folding get started? *Biochem. Sci.* **14**, 291–294.
- Barrick, D. & Baldwin, R. L. (1993a). Three-state analysis of sperm whale apomyoglobin unfolding. *Biochemistry*, **32**, 3790–3796.
- Barrick, D. & Baldwin, R. L. (1993b). The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Sci.* **2**, 869–876.
- Barrick, D., Hughson, F. M. & Baldwin, R. L. (1994). Molecular mechanism of acid denaturation. The role of histidine residues on the partial unfolding of apomyoglobin. *J. Mol. Biol.* **237**, 588–601.
- Chakrabartty, A., Schellman, J. A. & Baldwin, R. L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature*, **351**, 586–8.
- Chakrabartty, A., Kortemme, T. & Baldwin, R. L. (1994). Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**, 843–852.
- Creamer, T. P. & Rose, G. D. (1995). Interactions between hydrophobic side chains within  $\alpha$ -helices. *Protein Sci.* **4**, 1305–1314.
- Dobson, C. M. (1992). Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* **2**, 6–12.
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.
- Goto, Y. & Fink, A. L. (1990). Phase diagram for acidic conformational states of apomyoglobin. *J. Mol. Biol.* **214**, 803–805.
- Goto, Y., Calciano, L. J. & Fink, A. L. (1990a). Acid-induced folding of proteins. *Proc. Natl Acad. Sci. USA*, **87**, 573–577.
- Goto, Y., Takahashi, N. & Fink, A. L. (1990b). Mechanism of acid-induced folding of proteins. *Biochemistry*, **29**, 3480–3488.
- Greene, R. F. & Pace, C. N. (1974). Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme,  $\alpha$ -chymotrypsin and  $\beta$ -lactoglobulin. *J. Biol. Chem.* **17**, 5388–5393.
- Griko, Y. V. & Privalov, P. L. (1994). Thermodynamic puzzle of apomyoglobin unfolding. *J. Mol. Biol.* **235**, 1318–1325.
- Griko, Y. V., Privalov, P., Venyaminov, Y. S. & Kutysenko, V. P. (1988). Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol.* **202**, 127–138.
- Handel, T. & De Grado, W. F. (1990). *De novo* design of a  $Zn^{2+}$ -binding protein. *J. Am. Chem. Soc.* **112**, 6710–6711.
- Hughson, F. M. & Baldwin, R. L. (1989). Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates. *Biochemistry*, **28**, 4415–22.
- Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science*, **249**, 1544–1548.
- Hughson, F. M., Barrick, D. & Baldwin, R. L. (1991). Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry*, **30**, 4113–4118.
- Jennings, P. A. & Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science*, **262**, 892–896.
- Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M. & Hecht, H. H. (1993). Protein design by binary patterning of polar and nonpolar amino acids. *Science*, **262**, 1680–1685.
- Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F. & Goto, Y. (1995). Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering. *J. Mol. Biol.*, **249**, 215–228.
- Kuwajima, K. (1989). The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct. Funct. Genet.* **6**, 87–103.
- Landt, O., Grunert, H.-P. & Hahn, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene*, **96**, 125–128.
- Lin, L., Pinker, R. J., Forde, K., Rose, G. D. & Kallenbach, N. R. (1994). Molten globular characteristics of the native state of apomyoglobin. *Nature Struct. Biol.* **1**, 447–452.
- Loh, S. N., Kay, M. S. & Baldwin, R. L. (1995). Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl Acad. Sci. USA*, **92**, 5446–5450.
- Muñoz, V. & Serrano, L. (1994). Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct. Biol.* **1**, 399–409.
- Nakano, M. & Yang, J. T. (1981). Helical conformation of myoglobin and its cyanogen bromide-treated fragments in sodium dodecyl sulfate solution. *Arch. Biochem. Biophys.* **207**, 69–74.
- Peng, Z. & Kim, P. S. (1994). A protein dissection study of a molten globule. *Biochemistry*, **33**, 2136–2141.
- Ptitsyn, O. B. (1992). The molten globule state. In *Protein Folding* (Creighton, T. E., ed.), pp. 243–300, Freeman and Co., NY.
- Ptitsyn, O. B. & Uversky, U. N. (1994). The molten globule is a third thermodynamical state of protein molecules. *FEBS Letters*, **341**, 15–18.

- Raleigh, D. P. & De Grado, W. F. (1992). A *de novo* designed protein shows a thermally induced transition from a native to a molten globule-like state. *J. Am. Chem. Soc.* **114**, 10079–10081.
- Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethane sulfonyl  $\alpha$ -chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
- Schellman, J. A. (1978). Solvent denaturation. *Biopolymers*, **17**, 1305–1322.
- Schellman, J. A. (1990). A simple model for solvation in mixed solvents. Applications to the stabilization and destabilization of macromolecular structures. *Biophys. Chem.* **37**, 121–140.
- Scholtz, J. M., Barrick, D., York, E. J., Stewart, J. M. & Baldwin, R. L. (1995). Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc. Natl Acad. Sci. USA*, **92**, 185–189.
- Stryer, L. (1988). *Biochemistry*, W. H. Freeman and Company, NY.
- Teale, F. W. (1959). Cleavage of the haem-protein link by acidic methylethylketone. *Biochim. Biophys. Acta.* **35**, 543.
- Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J. & Wright, P. E. (1993). 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.

**Edited by P. E. Wright**

(Received 24 April 1995; accepted 15 June 1995)