J. Mol. Biol. (1994) 237, 588–601

Molecular Mechanisms of Acid Denaturation

The Role of Histidine Residues in the Partial Unfolding of Apomyoglobin

Doug Barrick†, Frederick M. Hughson‡ and Robert L. Baldwin§

Department of Biochemistry
Beckman Center
Stanford University Medical Center
Stanford, California 94305, U.S.A.

Apomyoglobin adopts a partly folded intermediate conformation (I), sometimes referred to as a molten globule intermediate, near pH 4. To determine which histidine residues trigger this partial unfolding reaction, we made mutants in which nine of the twelve histidine residues in the protein are substituted individually. We then measured acid and urea-induced unfolding curves for these substituted proteins. Two acid unfolding transitions are observed: native (N) to intermediate (I), and I to unfolded (U). These data were fitted using a simple three-state model which has been shown to give an adequate description of acid and urea-induced unfolding of wild-type apomyoglobin. The aim is to quantify changes in the apparent standard Gibbs energy differences between N, I and U, as well as the unfolding mechanism, that result from these substitutions, and to test how well the model fits data for substituted proteins. In most cases, the model fits the data reasonably well, and significant changes in fitted unfolding parameters of various mutants are also clearly visible in the primary data. The following conclusions are drawn. (1) Histidines 24 and 119 synergistically stabilize native apomyoglobin (N) at pH 8, but together destabilize N as pH is decreased below seven. (2) Histidine 36 stabilizes N when it is protonated. (3) Histidine substitutions in the heme-binding pocket (residues 64, 93 and 97) have little effect on the stability of N, suggesting that the heme-binding pocket is open. (4) Histidine substitutions affect the N to I transition but have little effect on the I to U transition. (5) The simple model we use to describe the unfolding of apomyoglobin cannot account for all the data, particularly the effects of the H36Q mutation. The effect of protonated histidine 36 on stabilizing N is not included in the model.

We suggest that breaking the hydrogen bond between histidines 24 and 119 by protonation when the pH is decreased from 6 to 4 is an important part of triggering the partial unfolding of N to I, and likewise that formation of the hydrogen bond between histidines 24 and 119 may be a rate-determining step in the kinetic process of forming N from I during refolding.

Keywords: protein folding; protein structure; protein stability; molten globule; apomyoglobin

1. Introduction

Apomyoglobin (apoMb‖, myoglobin without heme) adopts an intermediate (I) conformation at pH 4 that has been the focus of structural, thermodynamic, kinetic and theoretical studies (Breslow *et al.*, 1965; Kirby & Steiner, 1970; Shen & Hermans, 1972; Ptitsyn & Rashin, 1975; Kihara *et al.*, 1980;

† Present address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, U.S.A.

‡ Present address: Departments of Molecular Biology and Biochemistry, Harvard University, 7 Divinity Avenue, Cambridge, MA 02318, U.S.A.

§ Author to whom all correspondence should be addressed.

‖ Abbreviations used: apoMb, apomyoglobin; holoMb, holomyoglobin; CD, circular dichroism; $[\theta]_{222}$, mean molar residue ellipticity at 222 nm; N, I and U, the native, intermediate, and unfolded conformations of apomyoglobin; PCR, polymerase chain reaction; NOE, nuclear Overhauser effect; $K_{NI,app}$ and $K_{IU,app}$, apparent equilibrium constants between N and I and I and U respectively; $\Delta G^{\circ}_{NI,app}$ and $\Delta G^{\circ}_{IU,app}$, apparent Gibbs energy differences between N and I, and I and U respectively; n_{H1} and n_{H2}, the apparent number of proton binding sites linked to the N to I and I to U transitions respectively, $pK_{a1,I}$ and $pK_{a2,U}$, the pKa values of the n_{H1} and n_{H2} sites in I and U respectively; m_{NI} and m_{IU}, the urea sensitivities of the N to I and I to U reactions, respectively.

0022-2836/94/150588-14 $08.00/0

© 1994 Academic Press Limited

Irace *et al.*, 1981; Griko *et al.*, 1988; Hughson & Baldwin, 1989; Hughson *et al.*, 1990, 1991; Cocco & Lecomte, 1990; Cocco *et al.*, 1992; Brooks, 1992; Barrick & Baldwin, 1993*a*,*b*; Tirado-Rives & Jorgensen, 1993; Waltho *et al.*, 1993; Shin *et al.*, 1993*a*,*b*; Jennings & Wright, 1993). While the native (N) conformation is compact, and contains secondary and tertiary structure similar to that of holomyoglobin (holoMb, Griko *et al.*, 1988; Hughson *et al.*, 1990; Cocco & Lecomte, 1990), apoMb I shows decreased helix content (Kirby & Steiner, 1970; Griko *et al.*, 1988; Hughson *et al.*, 1990) and lacks the tight side-chain packing characteristic of globular protein cores (Hughson *et al.*, 1991). Recently, pulsed hydrogen exchange studies (Jennings & Wright, 1993) identified an early folding intermediate with a very similar pattern of amide NH protection to that seen for the equilibrium intermediate (Hughson *et al.*, 1990), suggesting the equilibrium intermediate is also a kinetic folding intermediate.

In a recent study, we measured quantitatively the pH and urea dependences of the relative stabilities of N, I, and the fully unfolded (U) conformations of apoMb (Barrick & Baldwin, 1993*a*). This analysis shows that the unfolding of N can be modeled as resulting from approximately two proton-binding sites with pK_a values around 7 in apoMb I and U, and which cannot bind protons (i.e. have abnormally low pK_a values) in N. Since histidine in unstructured polypeptide has a pK_a near 7 (Reynolds *et al.*, 1973; Anderson *et al.*, 1990; McNutt *et al.*, 1990), this pH dependence of unfolding may result from histidine residues.

Studies on other proteins show that conformational stability can be affected strongly by thermodynamic coupling to the acid-base equilibrium of histidine. For example, the stability of T4 lysozyme is dependent on pH in part because histidine 31 has an abnormally high pK_a (9·1) in the native protein (Anderson *et al.*, 1990). In a study of RNase T1, replacement of Glu58 with Ala has large effects on the pK_a values of two neighboring histidines and thus affects the pH dependence of stability (McNutt *et al.*, 1990). Replacement of a histidine with a glutamine at the C terminus of an α-helix in barnase has a pH-dependent effect on stability consistent with the abnormally high pK_a of the histidine (Sali *et al.*, 1988). In these studies, mutagenesis of either the histidines (Sali *et al.*, 1988; Anderson *et al.*, 1990) or residues interacting with the histidines (Anderson *et al.*, 1990; McNutt *et al.*, 1990) was critical in verifying the role of the histidine in pH-dependent stability changes.

Here we test whether the acid-induced transition from apoMb N to I can be attributed to a small number of histidine residues, as we have modeled (Barrick & Baldwin, 1993*a*), and if so, which histidines are responsible. We construct single mutants of nine of the twelve histidines in myoglobin, and also a double mutant involving histidines 24 and 119. Combinations of acid and urea are then used to analyze the effect of each mutation on the stabilities of N, I and U, and on the pH dependence of the stability of N. We fit our three-state stability equation (Barrick & Baldwin, 1993*a*) to the unfolding data in order to obtain quantitative unfolding parameters. Results of the denaturation studies here are consistent with a partial list of the histidine pK_a values in apoMb N (Cocco *et al.*, 1992) and with structural data for holo- and apoMb (Takano *et al.*, 1977; Cheng & Schoenborn, 1991; Cocco & Lecomte, 1990), and partially describe the molecular mechanism of acid-induced unfolding of apoMb N.

2. Materials and Methods

(a) *Choice of histidine mutations*

The 12 histidines in myoglobin are listed in Table 1, together with information regarding their environments in the crystal structure of holoMb, measured pK_a values of some histidines in apoMb N (Cocco *et al.*, 1992), and percent protonation at pH 5·7 from crystalline neutron diffraction studies of carbonmonoxymyoglobin (Cheng & Schoenborn, 1991). We are interested in replacing the histidines that have abnormal pK_a values in apoMb. For this reason, of the histidines for which pK_a values in apoMb N are known, we replaced all but histidine 116 (pK_a 6·6). For those histidines where pK_a values are not available, we calculated solvent-accessible surface area (see the legend to Table 1), and replaced histidines which have low solvent-accessible surface area in holoMb since burial may result in low pK_a in apoMb N. Based on calculated solvent-accessible surface areas, we mutated histidines 64, 82, 93 and 97. We note, however, that three of these four histidines contact heme in holoMb; thus their environments should be quite different in apoMb. Since histidines 24 and 119 are placed so as to interact with each other in holoMb and in apoMb (Takano, 1977; Cocco & Lecomte, 1990), we chose to replace them simultaneously as well as individually. For most mutations histidine is replaced with glutamine (Table 2) to preserve side-chain volume and, to a lesser degree, polarity. For histidines 24 and 119, we substitute valine and phenylalanine, respectively, since these two residues are frequently found together in globin sequences†.

The expression and purification system used here relies on the production of holoMb. We were initially concerned that mutation of histidine 93, which provides the proximal ligand to the heme iron (Kendrew, 1962), would be unable to bind heme; failure to bind heme may result in an apoMb that is unstable in *Escherichia coli*, and may seriously complicate purification. To circumvent this problem, we replaced histidine 93 with glycine, and then supplemented all growth and purification steps with 1 to 10 mM imidazole (see below). Our intention was to stabilize the H93G holoMb both *in vivo* and during purification by substituting imidazole as a proximal ligand in the cavity left by the mutation H93G (Eriksson *et al.*, 1992). Based on preliminary spectroscopic characterization, addition of imidazole produces an H93G holoMb conformation very similar to wild-type holoMb (D. B., unpublished results).

† In a list of 215 globin sequences, kindly provided by Donald Bashford, out of the 40 sequences containing valine at position 24, 30 contain phenylalanine at position 119.

Table 1
Histidines in myoglobin

Histidine†	Solvent exposure‡ (Å²)	Contacts in holoMb§	pK_a‖	Occupancy¶ (%)
12 (A10)	108·3	solvent		2
24 (B5)	2·8	E, G, H119 N^{ε}	<4·8	41
36 (C1)	46·5	G	8·1–8·2	0
48 (CD6)	106·1	solvent	5·2	48
64d (E7)	26·7	heme		0
81 (EF4)	135·9	solvent		85
82 (EF5)	4·8	N^{ε}2·6 Å from D141O^{γ}		100
93p (F8)	50·7	H, heme		N/D
97 (FG2)	79·9	heme, 4 H_2O, H93		97
113 (G14)	57·3	3 H_2O around N^{ε}	<5·5	14
116 (G17)	105·5	solvent	6·6	9
119 (GH1)	16·5	B, H24 N^{ε}	5·3–5·8	55

† Residue number, based on the sequence of wild-type sperm whale myoglobin. Myoglobin produced in the bacterial expression systems contains a formyl-methionine on the N terminus, however, as with protein derived from sperm whale, we take the following valine as residue number 1. Numbers in parenthesis indicate positions within helices and turns, and are as in Lesk & Chothia (1980).

‡ Determined by the method of Lee & Richards (1971) from the Brookhaven file 2MBN (Takano, 1977), with the heme omitted.

§ From Lesk & Chothia (1980), and from inspection of the crystal structure.

‖ pK_a values determined for apoMb at 25 °C (Cocco *et al.*, 1992).

¶ Degree of protonation determined from neutron diffraction studies of carbonmonoxymyoglobin at pH 5·7 (Cheng & Schoenborn, 1991).

(b) *Construction of mutants*

All DNA manipulations are performed essentially as described in Sambrook *et al.* (1989). All mutagenesis is performed on a derivative of a synthetic sperm-whale myoglobin gene, kindly provided by Drs Barry A. Springer and Steven G. Sligar. Mutagenesis of histidines 24 and 119 was achieved using the Kunkel method (Kunkel *et al.*, 1987). Histidine 113 was replaced using cassette mutagenesis. For other histidine mutations, cassette replacements are impractical due to a paucity of unique restriction sites. Instead we use a recombinant polymerase chain reaction (PCR) technique (Merino *et al.*, 1992; Higuchi, 1990). In this method, a single oligonucleotide is synthesized which spans the site to be mutagenized. The mutagenic oligonucleotide contains base substitutions coding for the desired substitution and is used in conjunction with a C-terminal primer in an initial round of PCR on pMb413b (a plasmid containing the wild-type myoglobin gene). The C-terminal primer is

Table 2
Best-fit parameters describing the acid-induced folding of apoMb histidine mutants

Mutant	n_{H1}†	$\Delta G^{\circ}_{NI,app}$‡ (kcal mol^{-1})	$\Delta G^{\circ}_{IU,app}$‡ (kcal mol^{-1})	√(Variance)§
H24V	2·26 (1·91–2·74)	4·01 (3·69–4·27)	2·08 (2·00–2·18)	657
H36Q	3·94 (N/D)	3·38 (N/D)	1·93 (N/D)	795
H48Q	2·40 (2·12–2·73)	4·04 (3·78–4·27)	1·95 (1·88–2·03)	703
H64Q	2·04 (1·93–2·13)	4·20 (4·07–4·33)	1·96 (1·88–2·04)	434
H82Q	2·61 (2·35–2·87)	4·70 (4·52–4·88)	1·86 (1·77–1·95)	595
H93G	2·71 (1·75–1·88)	4·60 (4·47–4·73)	2·05 (1·99–2·11)	518
H97Q	2·23 (2·05–2·43)	4·57 (4·37–4·75)	1·93 (1·84–2·02)	657
H113Q	2·27 (2·10–2·46)	4·19 (4·02–4·35)	2·16 (2·10–2·23)	469
H119F	2·58 (2·31–2·84)	3·84 (3·62–4·04)	2·09 (2·03–2·15)	520
H24V/H119F	1·24 (1·12–1·35)	5·38 (5·19–5·56)	2·38 (2·26–2·53)	684
Wild-type‖	2·25 (1·86–2·65)	4·55 (4·07–4·79)	2·06 (1·86–2·23)	1188

Parameters result from applying the constrained 3-state fitting procedure to the data set (see the text).

† Values in parentheses are 67% confidence limits determined during fitting. For H36Q, NonLin was unable to calculate 67% confidence limits, perhaps as a result of poor fitting.

‡ Apparent standard Gibbs energy differences are at 0 M urea, and pH ≫7, and are computed using the equation $\Delta G^{\circ}_{app} = -RT\ln K_{eq,app}$, where R is the gas constant ($1{\cdot}987 \times 10^{-3}$ kcal mol^{-1} K^{-1}), and T is 273·15 K. Uncertainties in free energy are calculated as $-RT[\partial(\ln K_{eq,app})/\partial K_{eq,app}]\sigma K_{eq,app}$, where $\sigma K_{eq,app}$ is the 67% confidence limit of K determined during fitting.

§ Square root of the variance in deg cm^2 dmol^{-1}.

‖ Parameters for wild-type unfolding are from Table I of Barrick & Baldwin (1993*a*), in which data from both experiments were combined (544 observations). Although fitting to experiment 1 of that paper alone produces a significantly lower variance (630 deg cm^2 dmol^{-1}) resulting in part from a major improvement in instrument optics, parameters are not significantly different between data sets.

homologous to a region of the plasmid just outside the coding sequence, and has on its 5′ end an extension (5′-AAGGCCTTGTCGACCCCC). The product of the primary PCR is treated with 3 units of T4-DNA polymerase (New England Biolabs, Beverly, MA) in 0·2 to 0·5 mM of all 4 deoxynucleotide triphosphates, and is purified by native polyacrylamide gel electrophoresis. The gel-purified product is mixed with 10 ng of pMb413b and subjected to 5 rounds of thermal cycling in the absence of primers. An N-terminal primer (reverse sequencing primer 1233, New England Biolabs) and a primer identical to the 5′ extension of the C-terminal primer listed above are added, and a 2nd round of PCR is carried out. The product is digested with the appropriate restriction enzymes, gel purified, and ligated into pMb413b. The T4-DNA polymerase treatment of the primary PCR product in an excess of dNTPs is designed to remove, *via* 3′→5′ exonuclease activity, the 3′ dA_2 dinucleotide that is added by Taq polymerase. Nevertheless, dideoxy-sequencing reveals that this step is sometimes unsuccessful; the extra nucleotides cause a frame-shift 5′ to the mutation. To avoid the frame-shift problem, mutagenic primers are currently designed so that their 5′ ends are immediately 3′ to the sequence TT. For all mutant constructs, the entire portion of the gene replaced in the ligation step was sequenced using the dideoxy method (Sanger, 1981) in order to avoid errors resulting from PCR.

(c) *Protein purification*

Mutant myoglobins are produced in the *E. coli* strain TB-1 as described by Barrick & Baldwin (1993*a*). Except for H93G, protein is purified from the supernatant of cleared lysates as described (Hughson *et al.*, 1991; Barrick & Baldwin, 1993*a*). For H93G, expression and purification are done in the presence of 10 mM imidazole·HCl at the appropriate pH values (1 mM imidazole is used in the cation exchange step in order to keep ionic strength low). This method results in high level expression, around 5 mg/l of cell culture. We cannot detect changes in chromatographic behavior or yield for any of the mutant proteins, although H64Q is bright red, rather than the usual reddish brown seen for wild-type protein. Significant spectral changes which correlate with the replacement of histidine 64 with glutamine have been noted (Shikama & Matsuoka, 1989). Proteins are typically greater than 95% pure as judged by Coomassie staining of SDS-polyacrylamide gels.

ApoMb is prepared by 2-butanone extraction of heme (Teale, 1959, Hughson *et al.*, 1991). Concentration of apoMb in stock solutions is determined using the method of Edelhoch (1967). ApoMb stocks are stored at 0°C to 4°C; all subsequent measurements are made within 48 h.

(d) *CD studies of apoMb mutants*

CD measurements are made as described in Barrick & Baldwin (1993*a*). Mean molar residue ellipticity at 222 nm ($[\theta]_{222}$), which measures α-helical structure, was monitored at 0°C as a function of urea concentration and pH. Typically, for each mutant, 3 acid-induced unfolding curves (in urea concentrations ranging from 0 to 3 M) and 4 urea-induced unfolding curves (at pH 8, 7, 6, and either 5, 4·5, or 5) are generated, giving approximately 160 separate CD measurements per mutant. We observe small but significant changes in ellipticity at 222 nm between different apoMb preparations from day to day; therefore all 7 unfolding curves for a given mutant are collected in one 24 h period. All samples were 1 μM in apoMb, and contained 2 mM total of sodium citrate plus citric acid. In a recent report describing apoMb precipitation as a function of urea concentration (De Young *et al.*, 1993) a solubility minimum was found at 2·4 M urea of around 2 mg apoMb/ml. In the current study, as in all previous circular dichroism studies from our laboratory, we use less than one-hundredth the concentration of this lower limit in solubility, in order to avoid the well-known problem of apoMb aggregation and precipitation (Rumen & Appella, 1962). For each measurement, 5 to 20 μl of apoMb stock is added to 2 ml of prechilled buffer containing the appropriate amount of urea. The sample is immediately mixed, transferred to a thermostatted (0°C) 1·00 cm quartz cuvette, and allowed to reach thermal and chemical equilibrium for 5 min. Ellipticity is then averaged over the next 120 s. The pH of each sample is determined at the conclusion of each titration with a Metrohm combined glass electrode (Model 6.0216.100) at room temperature.

(e) *Three-state fitting*

We apply a three-state model for acid and urea-induced unfolding to our mutant apoMbs. We use the program NonLin for Macintosh (Brenstein, 1991), to fit eqn (15) of Barrick & Baldwin (1993*a*) to the data. This eqn describes the relative stability of the 3 conformations using 2 reference equilibrium constants for the N to I and I to U reactions ($K_{NI,app}$ and $K_{IU,app}$; apparent reference conditions are high pH, no urea). In the model, the natural logarithm of these apparent equilibrium constants depends linearly on the urea molarity; the linear coefficients of change in $\ln(K_{NI,app})$ and $\ln(K_{IU,app})$ with molar urea are m_{NI}/RT and m_{IU}/RT, where R is the gas constant, and T is absolute temperature. This formulation is identical to the linear extrapolation model (Greene & Pace, 1974; Schellman, 1978; Santoro & Bolen, 1992). The influence of pH on $K_{NI,app}$ and $K_{IU,app}$ is modeled as the result of some apparent number of proton binding sites, n_{H1} and n_{H2}, which have low pK_a values in the N and I forms respectively, and normal pK_a values in I and U ($pK_{a1,I}$ and $pK_{a2,U}$ respectively Barrick & Baldwin, 1993*a*). In addition, $[\theta]_{222}$ values for N, I and U are fitted to the data; those for N and U are allowed to vary linearly with urea concentration, because the data are sufficient to evaluate the urea dependence for each case. The data for I are not sufficient to determine its urea dependence and $[\theta]_{222}$ for I is treated as a constant.

3. Results

(a) *Acid-induced unfolding of histidine mutants of apoMb*

Acid-induced unfolding of apoMb mutants in the absence of urea, monitored by CD at 222 nm, is shown in Figure 1. In each panel, wild-type protein is shown for comparison (open circles). Ellipticity in Figure 1 is scaled so that at high pH each apoMb has the same value of $[\theta]_{222}$. This manipulation allows denaturation curves to be compared, but has no effect on the position of unfolding midpoints or steepness of unfolding transitions.

Acid-induced unfolding of the single mutants H24V, H119F, the double mutant H24V/H119F, and wild-type apoMb show that the effects of single

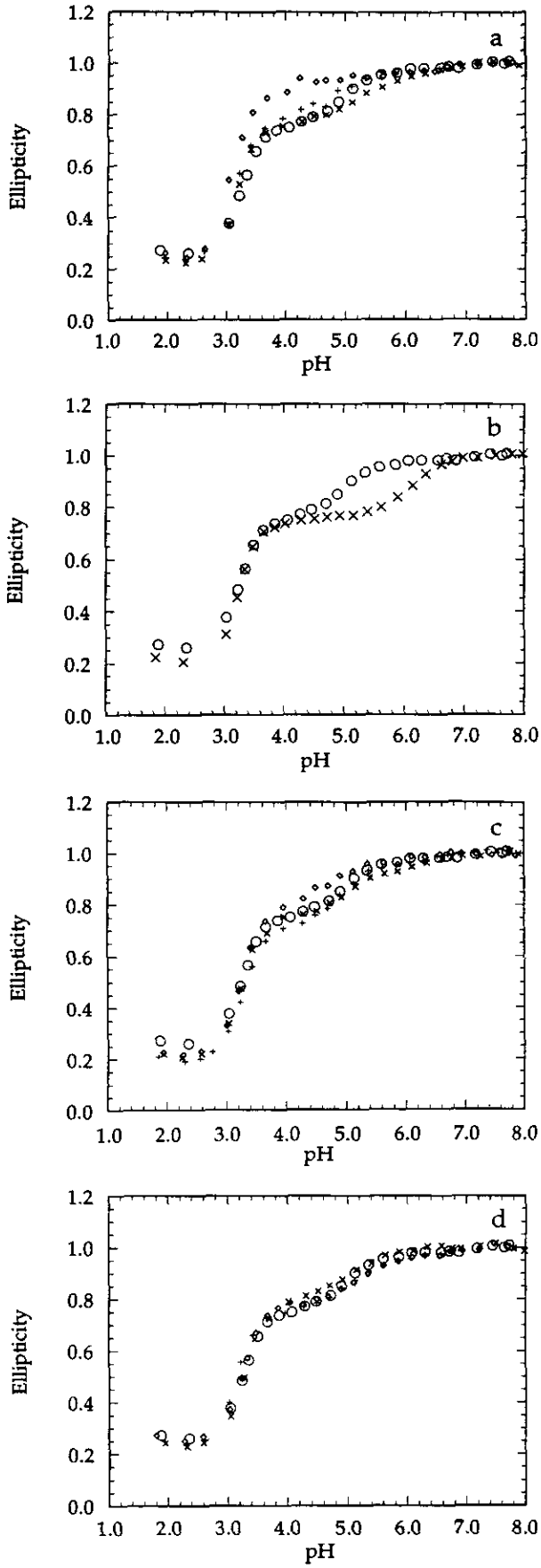

Figure 1. Acid-induced unfolding transitions of wild-type (○) and histidine mutants of apoMb in the absence of urea, monitored by CD at 222 nm. Data are scaled so that each apoMb has the same ellipticity at high pH. a: H24V (+); H119F (×); and H24V/H119F (◇). b: H36Q (×). c: H64Q (◇); H93G (×); H97Q (+). d: H48Q (◇); H82Q (×); and H113Q (+).

substitution are non-additive (Figure 1a). Although neither single mutant has large effects on either the N to I transition or the I to U transition, H24V/H119F unfolds at significantly lower pH. For wild-type, the midpoint for the N to I transition is around pH 5·0, and is well resolved from the low pH I to U transition. For H24V/H119F, unfolding appears to proceed in a single transition, with a pH midpoint of around 3·1 to 3·2. This increased resistance of the H24V/H119F N to acid demonstrates that the two replacements stabilize N relative to I at low pH. It is not clear from Figure 1a if the H24V/H119F unfolding transition results from a single two-state reaction from N to U, or if there are two transitions spaced too close to resolve the intermediate conformation. By adding one or two molar urea, the intermediate conformation is clearly resolved in acid unfolding (data not shown).

Acid-induced unfolding of H36Q shows a shift in the midpoint of the N to I transition from pH 5·0 to pH 6·1 (Figure 1b). This decreased resistance to acid demonstrates that N is destabilized relative to I by replacement of histidine 36. Although the N to I transition is strongly affected by this mutation, the I to U transition of H36Q is identical to wild-type.

The other histidine mutants in this study show little change from wild-type in their acid-induced unfolding curves. Acid-induced unfolding transitions of apoMb mutants where the replaced histidine is in the heme binding pocket (H64Q, H93G and H97Q) are very similar to wild-type (Figure 1c). There may be a small decrease in the pH midpoint of the N to I transition for H64Q; however, since the transition is not well defined, the decrease is probably not significant. For three remaining histidine mutants (H48Q, H82Q and H113Q) the acid-induced unfolding transitions are identical to wild-type (Figure 1d).

(b) *Urea-induced unfolding of histidine mutants of apoMb*

Urea-induced unfolding of apoMb mutants at pH 8, monitored by CD at 222 nm, is shown in Figure 2. Wild-type apoMb is shown for comparison (open circles). Ellipticity is scaled so that different apoMbs have the same value in the absence of urea. In urea-induced unfolding of wild-type at pH 8, unfolding occurs directly from N to U, with little I present at any urea concentration (Barrick & Baldwin, 1993*a*).

Urea-induced unfolding of the single mutants H24V, H119F, the double mutant H24V/H119F, and wild-type apoMb also show that the effects of single substitution are non-additive (Figure 2a). Urea midpoints for unfolding are decreased by 0·4 M for H24V and H119F, indicating a destabilization of

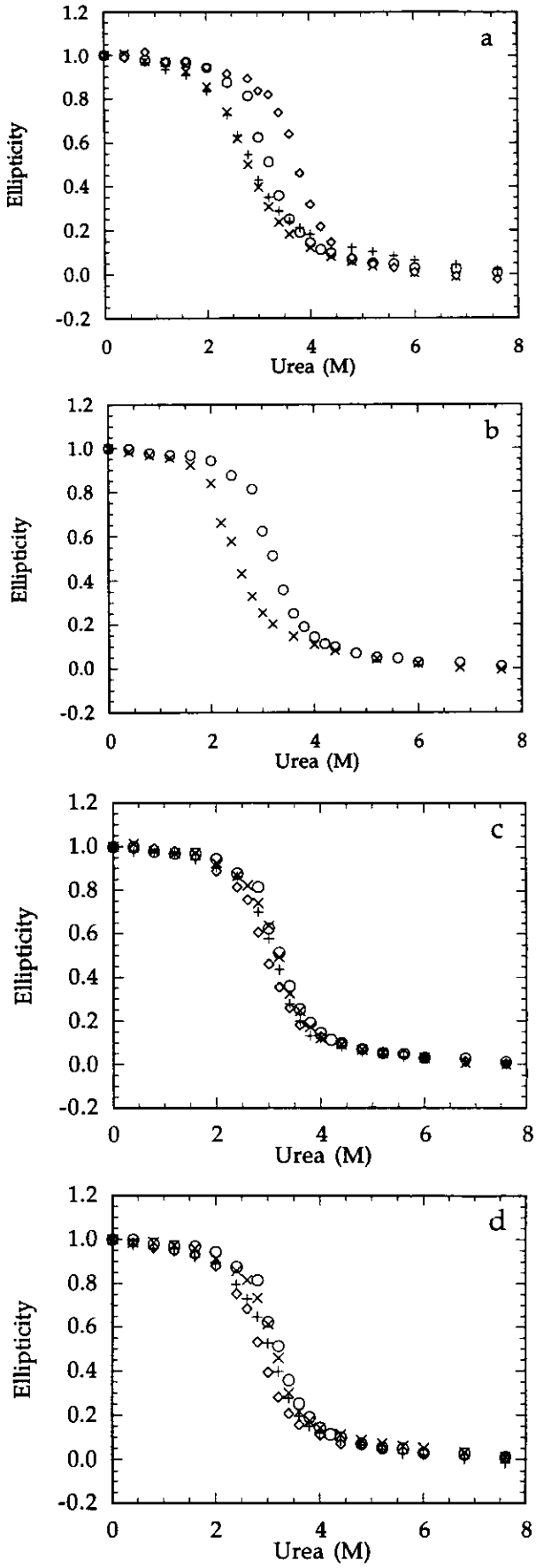

Figure 2. Urea-induced unfolding transitions of wild-type (○) and histidine mutants of apoMb at pH 8, monitored by CD at 222 nm. Data are scaled so that each apoMb has the same ellipticity at 0 M urea. a: H24V (+); H119F (×); and H24V/H119F (◇). b: H36Q (×). c: H64Q (◇); H93G (×); H97Q (+). d: H48Q (◇); H82Q (×); and H113Q (+).

N relative to U. For H24V/H119F, however, the urea midpoint is increased by 0·5 M, indicating N is stabilized relative to U. For H36Q, the urea midpoint for unfolding is decreased by 0·8 M (Figure 2b), indicating that N is destabilized relative to U.

For the remaining mutations, smaller changes are seen in urea-induced unfolding at pH 8 compared to wild-type. A small decrease in the urea midpoint of unfolding is seen for H64Q (0·3 M); for the other heme-pocket histidines (H93Q, H97Q) urea-induced unfolding is identical to wild-type (Figure 2c). H48Q shows a 0·4 M decrease in urea midpoint (Figure 2d); H82Q and H113Q are unchanged.

(c) *Three-state fitting of urea and acid-induced unfolding*

To quantify the changes in the apparent Gibbs energy differences between N, I and U resulting from mutation, and to identify changes in unfolding mechanism resulting from mutation, we apply a three-state analysis to each mutant. The details of this method are described elsewhere (Barrick & Baldwin, 1993*a*). The method relies on the use of a second denaturant, urea, to spread the acid-induced transitions over a wider range of pH. Such data are then fitted by a simple model which describes how the relative stabilities of N, I and U vary with pH and urea. We emphasize that this model contains several assumptions, as enumerated in Barrick & Baldwin (1993*a*), including an assumed two-state unfolding of I, and a very simplistic model for the pH dependence of the three-state equilibrium between N, I and U. Application of this model to the data presented here is justified on the grounds that it is the simplest physically reasonable model we can think of, and it adequately describes the acid and urea-induced unfolding of wild-type apoMb (Barrick & Baldwin, 1993*a*). Analysis of the data presented here using the model provides a quantitative measure of changes in the data that result from mutation. If the assumptions of the model are incorrect, the parameters determined in the fitting may not have exact thermodynamic or mechanistic meaning; nevertheless they are still useful as phenomenological quantities that describe the change in conformational stability and sensitivity of a given transition to hydrogen ion concentration.

For each mutant, we fit simultaneously all acid and urea-induced unfolding transitions. Initially we allowed all the parameters in the three-state model to be adjusted in the fitting procedure. Although in most cases this procedure gave parameters that accurately reproduced the data, we found that for six of the ten mutants the *m*-value for the N to I

transition showed a significant decrease to an average value of 0·6 (range 0·34 for H24V to 0·75 for H64Q) kcal mol^{-1} (M urea)$^{-1}$ compared with the wild-type value of 1·03 kcal mol^{-1} (M urea)$^{-1}$. Similar decreases in the urea sensitivity of the transition from native to intermediate conformations have been observed in a three-state analysis of the α-subunit of tryptophan synthase (Tsuji *et al.*, 1993). Although these and other authors (Shortle & Meeker, 1986) have interpreted changes in *m*-value as reporting on changes in native and denatured conformations respectively, an alternative explanation of the m_{NI} changes is the appearance of additional intermediates in the unfolding transitions (see the discussion of Figure 1b by Gittis *et al.* (1993) concerning changes in *m*-values caused by mutations in staphylococcal nuclease). Thus, the fitted values of m_{NI} are likely to have high associated uncertainties; errors in m_{NI} will be propagated into $\Delta G^{\circ}_{NI,app}$ and n_{H1}, the parameters of primary interest here. Therefore, in this paper we apply a constrained three-state fitting procedure in which those parameters expected to remain constant upon histidine replacement are fixed at wild-type values. These parameters are m_{NI} and m_{IU} (1·03 and 1·01 kcal mol^{-1} (M urea)$^{-1}$, pK_{a1} (7·05), pK_{a2} (4·59), and n_{H2} (1·31 proton-binding sites), and were obtained reproducibly using different data sets for wild-type apoMb (Barrick & Baldwin, 1993*a*). The adjustable parameters in the constrained fitting procedure are n_{H1}, $K_{NI,app}$, and $K_{IU,app}$, and also the helix contents of N, I and U. This constrained fitting procedure simplifies the interpretation of the effects of histidine replacement, and allows direct comparison of the parameters of interest. Furthermore, the use of this constrained fitting method produces values consistent with the primary data, with what is known about the structure, and with a subset of measured apoMb histidine pK_a values. It should be kept in mind, however, that in some cases the constraints we apply may provide a poor description of urea-induced unfolding.

Fitted unfolding parameters and square roots of variances (standard deviations) for the ten mutants are listed in Table 2. Standard deviations range from 795 to 434 deg cm^2 $dmol^{-1}$. To illustrate the quality of the fits, we show the data set with the best fit (i.e. lowest variance, H64Q; Figure 3a,b), median fit (H82Q; Figure 3c,d), and worst fit (highest variance, H36Q; Figure 3e,f). In each case, urea-induced unfolding curves are reasonably well reproduced using the fitting parameters (Figure 3b,d,f). In acid unfolding curves (Figure 3a,c,e), discrepancies are seen between data and fitted curves at low urea concentrations, pH 4 to 5·5. For H64Q these discrepancies are quite small, but for H36Q they are very large. The fit to the H82Q data set (that which produced a median variance) is satisfactory for our analysis.

The most significant changes in $\Delta G^{\circ}_{NI,app}$ are an increase of around 0·8 kcal mol^{-1} for the double mutant H24V/H119F and a decrease of around 1·2 kcal mol^{-1} for H36Q (Figure 4a, Table 2); the

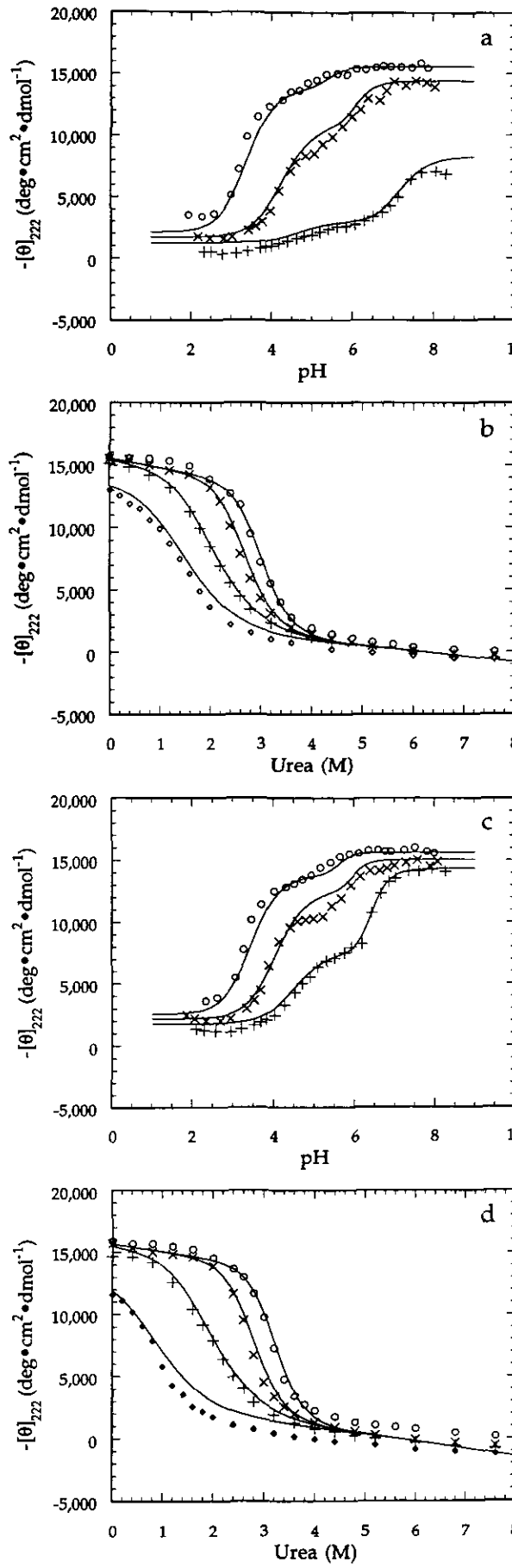

Fig. 3.

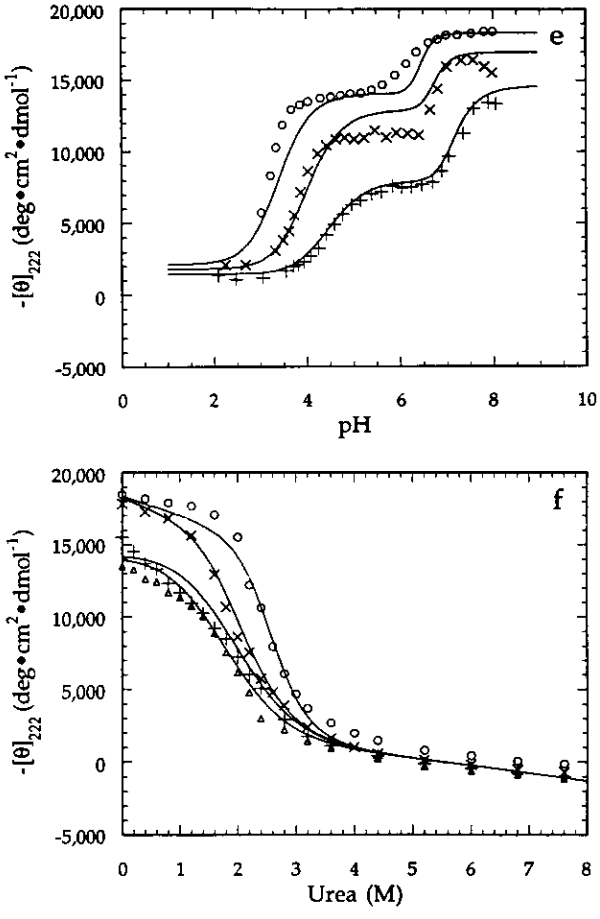

Figure 3. Three-state analysis of acid- and urea-induced unfolding of representative apoMb mutants. H64Q; (a), (b): H82Q; (c), (d): and H36Q; (e), (f) give the lowest, median, and highest variances in the three-state fitting procedure. Panel a: acid-induced unfolding in 0 (○); 1·5 (×); and 3 (+) M urea. Panels c and e: acid-induced unfolding in 0 (○); 1 (×); and 2 (+) M urea. Panels b, d and f: urea-induced unfolding at pH 8 (○); 7 (×); 6 (+); 5 (Δ); 4·5 (◇); and 4 (◆).

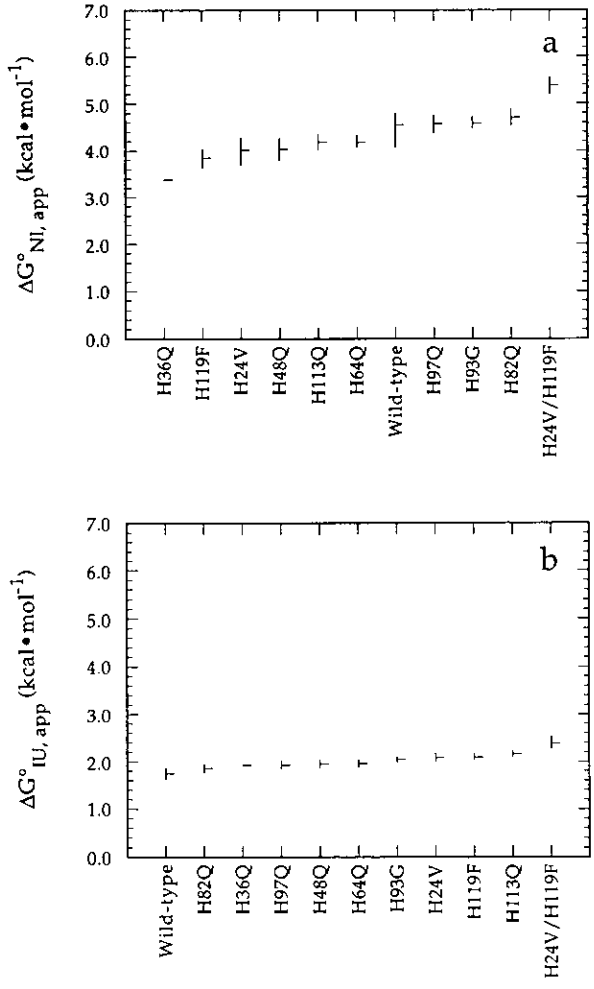

Figure 4. Standard apparent Gibbs energy differences between: a, the N and I; b, I and U conformations for wild-type and histidine mutants of apoMb, determined from the constrained 3-state fitting. Apparent energy differences are in the absence of urea, at high pH. The vertical bars show the 67% confidence intervals determined from the fitting, the horizontal ticks show the best estimates. Mutants are arranged in order of increasing conformational stability. The confidence limits for H36Q are not known.

H36Q value is more uncertain because of poor fitting (Figure 3e). Decreases in $\Delta G°_{NI,app}$ of 0·7 and 0·5 kcal mol^{-1} are seen for H119F and H24V respectively; the latter value is within the 67 percent confidence limits of the wild-type value. Fitting of all other mutants gives smaller, insignificant changes in $\Delta G°_{NI,app}$.

$\Delta G°_{IU,app}$ is less sensitive to mutation (Figure 4b) than $\Delta G°_{NI,app}$. The largest change in $\Delta G°_{IU,app}$ compared with wild-type is an increase of 0·3 kcal mol^{-1} for H24V/H119F. The statistical significance of such a small change is slight; indeed the 67 percent confidence limits of $\Delta G°_{IU,app}$ for H24V/H119F and wild-type nearly overlap (Table 2).

The apparent number of titratable groups linked to the N to I transition (n_{H1}), as determined by the fitting procedure, is not significantly different from wild-type for most mutations (Figure 5). For the double mutant, however, n_{H1} is 1·02 titratable groups lower than for wild-type. For H36Q, n_{H1} is 1·75 titratable groups higher than for wild-type, but the confidence intervals are uncertain because of poor fitting.

4. Discussion

We studied the urea and acid-induced unfolding of a set of mutants in which histidine residues are replaced with residues that do not titrate. In an earlier study we reported qualitative analysis of the effect of packing mutations on the acid unfolding transitions of apoMb (Hughson *et al.*, 1991). In the current study, we apply a three-state model (Barrick & Baldwin, 1993*a*), in order to quantify the effects of mutations and also to test the model. Since the fitting procedure relies heavily on compu-

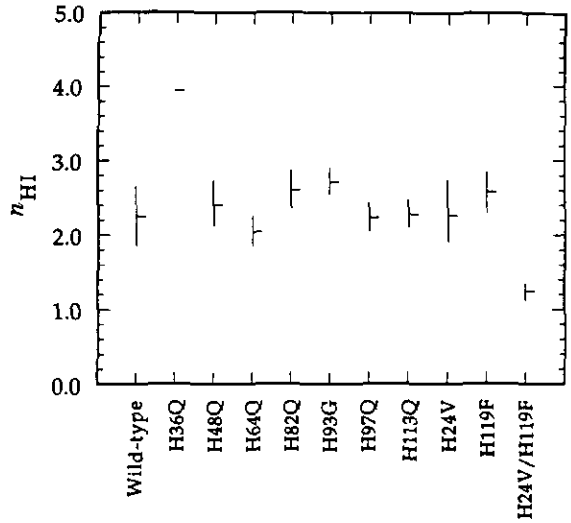

Figure 5. Apparent number of protons linked to the N to I transition (n_{H1}) for wild-type and histidine mutants of apoMb, determined from the constrained 3-state fitting. The vertical bars show the 67% confidence intervals determined from the fitting, the horizontal ticks show the best estimates.

tational methods, which put considerable distance between the fitted parameters (and associated conclusions) and the primary data, we begin our discussion by showing that significant changes in the fitting parameters are correlated with obvious changes in the primary data. We then discuss how these changes relate to the structure and stability of wild-type myoglobin. In doing so we compare fitted parameters, but the reader should keep in mind that the same conclusions can be made qualitatively from the primary data.

(a) *Comparison of fitting parameters with the primary data*

The largest changes in fitted $\Delta G^{\circ}_{NI,app}$ from wild-type are for H24V/H119F (0·8 kcal mol^{-1} increase in stability of N relative to I) and H36Q (1·2 kcal mol^{-1} decrease in stability). These changes are reflected in altered pH midpoints for unfolding: the N to I transition of H24V/H119F is shifted by 1·5 to 2·0 units to lower pH (Figure 1a), while that of H-36Q is shifted 1·0 unit to higher pH (Figure 1b). These effects are also seen in urea-induced unfolding at pH 8: the unfolding midpoint for H24V/H119F is increased by 0·5 M urea (Figure 2a), while that of H36Q is decreased by 0·8 M urea (Figure 2b). H119F shows a decrease in $\Delta G^{\circ}_{NI,app}$ compared with wild-type apoMb of 0·7 kcal mol^{-1}. Consistent with this modest decrease in stability is a small shift of the N to I transition to higher pH in acid-induced unfolding, and a decrease of the urea-induced unfolding midpoint by 0·5 M. For the other mutant apoMbs, changes in $\Delta G^{\circ}_{NI,app}$ are not significant; this behavior is also seen in the primary data (Figures 1c,d; 2c,d).

A change in n_{H1}, the apparent number of proton-binding sites linked to the N to I transition, is harder to visualize in the primary data since determination of n_{H1} requires the calculation of the first derivative of the acid-unfolding curve with respect to pH†. Nevertheless, for H36Q, where the fitted value of n_{H1} increases from wild-type by 1·75 groups, increased cooperativity can be seen both in pH titrations at 1 M urea (compare Figure 3c and e) and in the decrease in stability to urea with decreasing pH (compare urea-induced unfolding at pH 6 through 8 of Figure 3f with Figure 3b,d and also with the wild-type data in Figure 2 of Barrick & Baldwin, 1993*a*).

(b) *H24V, H119F and H24V/H119F*

A comparison of these three mutants and wild-type shows synergism between histidines 24 and 119, both in $\Delta G^{\circ}_{NI,app}$ and n_{H1}. While $\Delta G^{\circ}_{NI,app}$ is decreased for each single mutant, it is increased by 0·8 kcal mol^{-1} for the double mutant (Table 2). Similarly, n_{H1} for the double mutant, but not for either single mutant, drops substantially (from 2·25 protons for wild-type to 1·24 protons for the double mutant).

These results can be rationalized based on the structure and pK_a values of holoMb and apoMb. In holoMb, histidines 24 and 119 are buried in the hydrophobic core (Table 1). Stabilization of N relative to I and U in H24V/H119F is likely to result from removing two polar side-chains from the hydrophobic core, and replacing them with the hydrophobic side-chains of valine and phenylalanine, which occupy roughly the same volume. In holoMb, histidines 24 and 119 are nearly coplanar and hydrogen bonded to each other through their N$^{\varepsilon}$ atoms (2·5 Å separation; Takano, 1977). Histidine 24 also appears to form a hydrogen bond between its N$^{\delta}$ and the peptide carbonyl oxygen of aspartate 20 (Takano, 1977; Cheng & Schoenborn, 1991)‡. We propose that the hydrogen bond between histidines 24 and 119 stabilizes N; the decreased stability of each single mutant is expected, since replacement breaks this hydrogen bond and buries unsatisfied hydrogen-bond function. Using mutant-cycle formalism (Horovitz & Fersht, 1990), the interaction energy between residues at positions 24 and 119 is 2·1 kcal mol^{-1} at high pH (Figure 7).

† Although n_{H1} does have a direct influence on the position of the pH midpoint of the N to I transition, such displacements cannot be distinguished from changes in stability which result from pH independent interactions such as van der Waals contacts and changes in hydrophobicity.

‡ The structure surrounding histidines 24 and 119 in apoMb is likely to be similar to that in holoMb for the following reasons. First, the pK_a values and chemical shifts of histidines 24 and 119 do not change when heme is removed (Cocco *et al.*, 1992). Second, intraresidue NOEs are observed in apoMb between histidines 24 and 119, and also with other nearby residues in the holoprotein (Cocco & Lecomte, 1990).

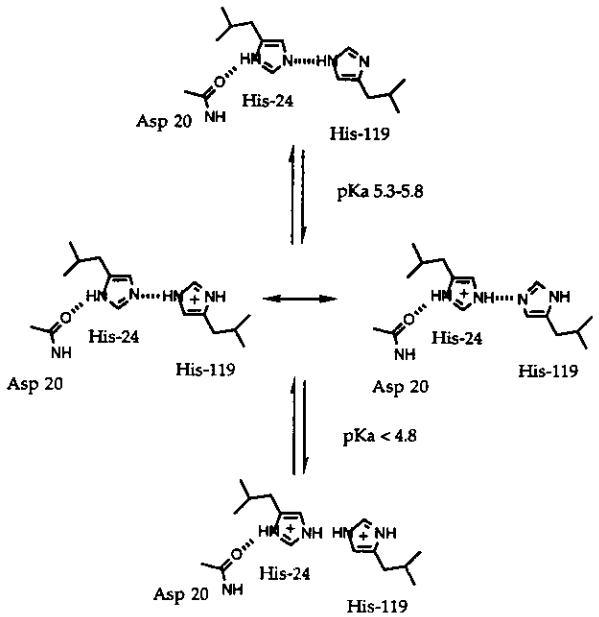

Figure 6. Proposed protonation states of histidines 24 and 119 in native apoMb. pK_a values are from Cocco *et al.* (1992). Evidence for a hydrogen bond between histidines 24 and 119 comes from neutron diffraction studies of Cheng & Schoenborn (1991) and from Dalvit & Wright (1987). In the fully deprotonated form, we have drawn His24 as the H-3 tautomer; this tautomer is only slightly less stable (Reynolds *et al.*, 1973). We propose that in the singly protonated form, the positive charge is delocalized across the two imidazole rings as shown.

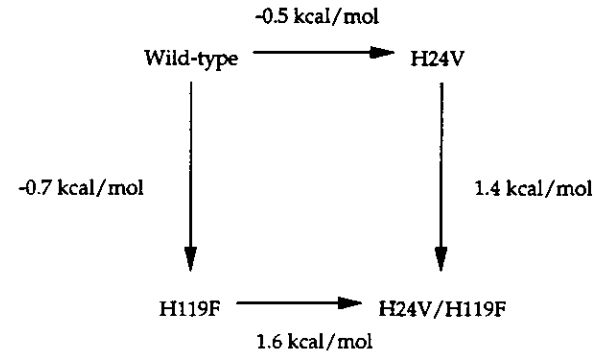

Figure 7. Mutant-cycle for H24V/H119F at pH 9. Values of $\Delta\Delta G^{\circ}_{NI,app}$, the difference in apparent fitted standard unfolding Gibbs energies of N to I, are shown for each mutation. Arrows determine the sign of $\Delta\Delta G^{\circ}_{NI,app}$; for instance, H24V minus wild-type is given on top. Since Gibbs energies are taken in the direction of unfolding, positive values for $\Delta\Delta G^{\circ}_{NI,app}$ indicate that the mutation is stabilizing to N. For double mutants in which the 2 mutated residues do not interact energetically, $\Delta\Delta G^{\circ}_{app}$ values on opposite sides of the cycle should be equal. The apparent Gibbs energy of interaction is taken as the difference in energy of one mutation in the absence *versus* presence of the other. Here, this apparent interaction energy is 1·6 minus −0·5, or 2·1 kcal mol^{-1}.

Although it is not possible to attribute this interaction energy to the two histidines alone†, the sign and magnitude of this energy are consistent with formation of a stable hydrogen bond.

The decrease in the fitted value of n_{H1} by one proton for H24V/H119F is consistent with the high energy penalty which must occur when these adjacent histidines are protonated in the low dielectric core of N. Presumably it is this penalty which shifts the pK_a of histidine 24 below 4·8 (Cocco *et al.*, 1992). The observation that neither H24V nor H119F shows any decrease, compared with wild-type, in the fitted value of n_{H1}, may be understood by considering the proposed protonation scheme shown in Figure 6‡. In the first protonation step in wild-type (pK_a 5·3 to 5·8), we propose that the destabilization associated with burial of a positive charge in the core of the protein is mitigated by charge delocalization across the two imidazole rings (Figure 6). When, however, one or the other imidazole group is replaced with a propyl or phenyl group (H24V, H119F), this delocalization is no longer possible. We hypothesize that this first (and only) protonation step becomes more destabilizing in either single mutant (lower pK_a), and effectively replaces the pH-dependent contribution of the second protonation step ($pK_a < 4·8$) seen in wild-type apoMb N. This prediction can be tested by measuring the pK_a of histidine 119 in H24V and that of histidine 24 in H119F.

The results and interpretation above for histidines 24 and 119 are consistent with data regarding the structure of apoMb I. Hydrogen exchange studies indicate that helices A, G and H are formed in apoMb I, and suggest that the intervening helices B through E are unfolded (Hughson *et al.*, 1990). Examination of the holoMb crystal structure (Takano, 1977) shows that the helices A, G and H, and the helices B through E form two separate subdomains (Hughson *et al.*, 1990, 1991). Histidines 24 and 119 are at this subdomain interface in holoMb and presumably in apoMb N (Cocco *et al.*, 1992; Cocco & Lecomte, 1990). Dissociation of these subdomains (followed by loss of packing for AGH (Hughson *et al.*, 1991), and more extensive unfolding for B-E) would relieve the unfavorable interaction which develops between His24 and His119 at low pH. In other words, what is known about the structures of N and I is consistent with

† The interaction energy could equally well result from stabilizing interactions between valine 24 and phenylalanine 119, or from repulsions between histidine 24 and phenylalanine 119 or between valine 24 and histidine 119.

‡ Although the structure around histidines 24 and 119 in wild-type apoMb probably resembles that in holoMb, at present we know little about the structures of the mutant apoMbs. NMR studies to examine the structures of the mutants is still under way: preliminary results suggest that much of the native structure seen in wild-type protein is retained in the mutants (M. Cocco & J. Lecomte, work in progress). Until a complete picture of these mutant structures is in hand, the structural interpretation given below should be taken as speculative.

both the pK_a values of His24 and His119, and the stability changes resulting from their replacement.

Our results, taken together with our analysis of the titration behavior of histidines 24 and 119, suggests that protonating this hydrogen-bonded pair is a major factor in triggering the partial unfolding of N to I as the pH is lowered from 6 to 4. Jennings & Wright (1993) have shown that, when urea-unfolded apoMb is allowed to refold at pH 6, the intermediate that is formed within the stopped-flow mixing time has the same protected amide protons as the equilibrium pH 4 folding intermediate. We suggest that formation of the hydrogen bond between histidines 24 and 119 may be a rate-determining step in the refolding to N of this intermediate.

(c) *H36Q*

Analysis of the data given here suggests that histidine 36 stabilizes apoMb N. For H36Q, the midpoint for urea-induced unfolding is decreased, the midpoint of the N to I transition in acid-induced unfolding is increased, and $\Delta G^\circ_{NI,app}$ is decreased by 1·2 kcal mol^{-1} compared to wild-type. Furthermore, the pH dependence of the N to I transition is increased: the acid-induced unfolding transition appears to be steeper, especially around pH 7 (Figure 3e, 1 M urea (×)), and n_{H1} is increased by 1·75 protons. Based on both the crystal structure of holoMb and the pK_a of histidine 36 (Cocco *et al.*, 1992), these effects can be partially rationalized. In the holoMb crystal structure, the side-chain of histidine 36 is packed against the side-chain of aspartate 38; these two residues may stabilize the protein through formation of a salt bridge. Additionally histidine 36 packs against phenylalanine 106. A stabilizing interaction between phenylalanine and histidine has been observed in a peptide helix; this interaction is strongest when the histidine is protonated (Armstrong *et al.*, 1993). Stabilization of apoMb N by charged histidine 36 should lead to an increase in pK_a; indeed, the pK_a of histidine 36 is 8·2 in both holo- and apoMb N (Cocco *et al.*, 1992). The observation that histidine 36 has the same value in holo- and apoMb N, together with the retention of an NOE between histidine 36 and phenylalanine 106 in apoMb N, indicates that the structure around histidine 36 is retained in apoMb (Cocco *et al.*, 1992). Based on the above observations, the destabilization of apoMb N displayed by H36Q is expected, as the replacement eliminates two potentially stabilizing interactions. Unlike the groups that are responsible for unfolding N at acidic pH, histidine 36 should stabilize the protein when the imidazole ring is protonated. As a result, elimination of histidine 36 should increase the sensitivity of the apoMb N stability to pH between its pK_a in N (8·2) and its pK_a in the I and U conformations (presumably 6·6 to 7), as is observed (Figure 3e, 1 M urea (×); Table 2).

The argument above also provides an explanation for the poor fitting of the H36Q data. The model on which the fitting is based describes the acid-unfolding of N as resulting from groups with abnormally low pK_a values in N. Removal of a group whose pK_a is abnormally high in N (a pK_a shift not accounted for by our simple model) may be expected to produce a pH dependence of the N to I transition that is inconsistent with our simple model. Stabilization of apoMb N at low pH by His36 shows that the model is oversimplified.

(d) *Heme binding pocket mutations*

The mutations H64Q, H93G, and H97Q show little change in either acid and urea-induced unfolding transitions, or in fitting parameters. In the crystal structure of holoMb, histidines 64, 93 and 97 contact the heme. The identical stability at all pHs for these mutant and wild-type apoMbs indicates that the heme pocket remains open; that is, removal of heme does not cause a collapse to a tightly packed structure. This is consistent with the fast binding of heme to apoMb (Antonini & Brunori, 1971; Shen & Hermans, 1972; Kawamura-Konishi *et al.*, 1988), two-dimensional NMR studies of the apoMb N (Cocco *et al.*, 1992), and NMR studies of dyes bound to the heme pocket of apoMb (Daniel Pierce & Steve Boxer, personal communication).

(e) *Comparison of* $\Delta G^\circ_{NI,app}$ *and* $\Delta G^\circ_{IU,app}$

Application of the three-state model to the different mutant proteins allows us to compare changes in apparent stability for the three conformations of apoMb. Changes in $\Delta G^\circ_{NI,app}$ are larger than in $\Delta G^\circ_{IU,app}$. In Figure 8, $\Delta G^\circ_{IU,app}$ is

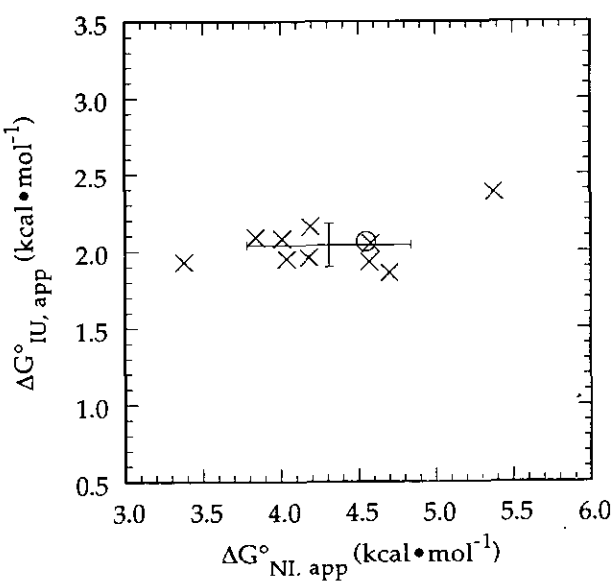

Figure 8. Comparison of apparent standard Gibbs energy differences between N and I with those between the I and U conformations for wild-type (○) and histidine mutants (×) of apoMb. The same scale is used for both axes, so that the spread of the data in each dimension can be compared; the crossbars show the root-mean-square deviation in each dimension.

plotted against $\Delta G^{\circ}_{NI,app}$ for each mutant. The scale is identical on each axis, so that the ranges of $\Delta G^{\circ}_{NI,app}$ and $\Delta G^{\circ}_{IU,app}$ can be compared. Crossbars show the calculated root-mean-square deviation of the distribution in each dimension. Although statistically crude, these root-mean-square deviations provide an analytical measure of the spread of the data; comparison of the root-mean-square-deviation in $\Delta G^{\circ}_{NI,app}$ with that in $\Delta G^{\circ}_{IU,app}$ shows that the mutations studied here affect the stability of N relative to I about four times more than that of I relative to U. If the stability of U is not affected by mutation, the factor of four above should reflect the sensitivity to mutation of N compared with I. Although these energy differences are apparent quantities, and are likely to suffer from inadequacies of the model (such as those implicated from the poor fitting of H36Q, and from potential multi-state unfolding of I), we point out again that the primary data support the insensitivity of I to mutation (Figure 1).

A similar conclusion was reached in a qualitative study of helix packing mutants in apoMb (Hughson *et al.*, 1991; Barrick & Baldwin, 1993*b*), and was taken as evidence that close packing interactions between helices are lost in the N to I transition. An alternative explanation, based on molecular dynamics simulations of apoMb in water, has been proposed by Tirado-Rives & Jorgensen (1993) in which close helix packing is retained in apoMb I, but does not involve the residues that were replaced in the study of Hughson *et al.* (1991). It will be interesting to see if any of the nine residues studied here are involved in the proposed apoMb I helix–packing interactions of Tirado–Rives & Jorgensen.

5. Conclusions

In the present study we replaced nine different histidine residues and studied the effects of these replacements on the stability of sperm whale apoMb. Our motivations were: (1) to learn about the molecular mechanisms that control acid unfolding of apoMb N; (2) to examine quantitatively the effects of histidine mutations on the relative stabilities of apoMb N, I and U; (3) to test the generality of a simple three-state model for apoMb unfolding (Barrick & Baldwin, 1993*a*).

We draw the following conclusions regarding the molecular mechanisms that control acid unfolding of apoMb N. Histidines 24 and 119, which are H-bonded to each other in holoMb, account for a major part of the pH dependence of the N to I transition: as the pH is decreased below 7, histidines 24 and 119 destabilize apoMb N. The double mutant H24V/H119F is strongly stabilized against acid unfolding and the apparent number of protons taken up in the N to I transition at low pH is reduced from 2·2 in wild-type to 1·2 in the double mutant. This dependence cannot, however, be attributed to either imidazole group alone. Rather, the interaction between these two groups creates unique proton affinities, such that the two imidazole groups must be considered as a whole. Histidine 36, which has a high pK_a (8·2) in N and in the crystal structure is close to two groups that might stabilize the protonated form of the side-chain, appears to influence the pH dependence of the stability of N in the opposite way: when it is protonated, histidine 36 stabilizes apoMb N. The mutant H36Q is destabilized against acid unfolding. The apparent number of protons taken up in the N to I transition at low pH is increased to 3·9, but this value is questionable because our model is unable to fit the H36Q unfolding transitions (see below). We fail to detect any change in the pH dependence of the stability of N in any of the other histidine mutants studied here, suggesting that the remaining pH dependence of 1·2 protons for the N to I reaction in H24V/H119F probably results from small pK_a shifts of several groups.

We draw the following conclusions regarding the generality of the three-state model for apoMb unfolding. Most of the mutant data obtained is fitted reasonably well. The origins of all significant changes in fitted unfolding parameters can be seen clearly in the primary data. This consistency between the primary data and the fitting suggests that the three-state model can be used to analyze data for mutants. Nonetheless, the results of fitting for some mutants are unimpressive; the most obvious example is H36Q, whose acid unfolding behavior is poorly fitted by our model. Given the data presented here for H36Q, the results of the fitting, and other structural data (Takano, 1977, Cocco *et al.*, 1992), we conclude that the molecular mechanism which causes N to unfold at low pH is more complicated than our original model (Barrick & Baldwin, 1993*a*). The increase in n_{H1} found for H36Q shows that values of n_{H1} obtained by using our model must be treated with caution. Although useful in extrapolating observed folding equilibria with pH, values of n_{H1} and $pK_{a1,I}$ found using our model do not have simple, direct correspondence to structural features or mechanisms; rather they should be taken as apparent quantities describing the macroscopic properties of apoMb.

We are very grateful to Barry Springer and Steve Sligar for giving us the sperm whale myoglobin gene. We thank Valerie Dagget for motivating the initial phase of this project. We thank Melanie Cocco and Juliette Lecomte for communicating results prior to publication, and also for helpful discussions of the results. We thank Donald Bashford for providing an extensive list of aligned globin sequences, Thomas Kiefhaber for useful suggestions regarding recombinant PCR and primer design, and Robert Brenstein for useful discussion of nonlinear least-squares fitting to the data. D.B. acknowledges the support of a Howard Hughes Medical Institute predoctoral fellowship. This work was supported by NIH grant GM19988.

References

Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990). pH-Induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy

of folding of T4 Lysozyme. *Biochemistry*, **29**, 2403–2408.

Antonini, E. & Brunori, M. (1971). *Hemoglobin and Myoglobin in Their Reaction with Ligands*. North-Holland Publishing Company, Amsterdam.

Armstrong, K. M., Fairman, R. & Baldwin, R. L. (1993). The (i, $i+4$) Phe-His interaction studied in an alanine-based α-helix. *J. Mol. Biol.* **230**, 284–291.

Barrick, D. & Baldwin, R. L. (1993*a*). Three-state analysis of sperm whale apomyoglobin folding. *Biochemistry*, **32**, 3790–3796.

Barrick, D. & Baldwin, R. L. (1993*b*). The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Science*, **2**, 869–893.

Brenstein, R. J. (1991). NonLin for Macintosh, Robelko Software, Carbondale, IL.

Breslow, E., Sherman, B., Hardman, K. D. & Gurd, F. R. N. (1965). Relative conformations of sperm whale metmyoglobin and apomyoglobin in solution. *J. Biol. Chem.* **240**, 304–309.

Brooks, C. L., III (1992). Characterization of "native" apomyoglobin by molecular dynamics simulation. *J. Mol. Biol.* **227**, 375–380.

Cheng, X. & Schoenborn, B. P. (1991). Neutron diffraction study of carbonmonoxymyoglobin. *J. Mol. Biol.* **220**, 381–399.

Cocco, M. J. & Lecomte, J. T. J. (1990). Characterization of hydrophobic cores in apomyoglobin: a proton NMR spectroscopy study. *Biochemistry*, **29**, 11067–11072.

Cocco, M. J., Kao, Y.-H., Phillips, A. T. & Lecomte, T. T. J. (1992). Structural comparison of apomyoglobin and metaquomyoglobin: pH titrations of histidines by NMR spectroscopy. *Biochemistry*, **31**, 6481–6491.

Dalvit, C. & Wright, P. E. (1987). Assignment of resonances in the ^{1}H nuclear magnetic resonance spectrum of the carbon monoxide complex of sperm whale myoglobin by phase-sensitive two-dimensional techniques. *J. Mol. Biol.* **194**, 313–327.

De Young, L. R., Dill, K. A. & Fink, A. L. (1993). Aggregation and denaturation of apomyoglobin in aqueous urea solutions. *Biochemistry*, **32**, 3877–3886.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.

Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1992). A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature (London)*, **355**, 371–373.

Gittis, A. G., Stites, W. E. & Lattman, E. E. (1993). The phase transition between a compact denatured state and a random coil state in staphylococcal nuclease is first-order. *J. Mol. Biol.* **232**, 718–724.

Greene, R. F. J. & Pace, C. N. (1974). Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin, and β-lactoglobulin. *J. Biol. Chem.* **249**, 5388–5393.

Griko, Y. V., Privalov, P. L., Venyaminov, S. Y. & Kutyshenko, V. P. (1988). Thermodynamic study of apomyoglobin structure. *J. Mol. Biol.* **202**, 127–138.

Higuchi, R. (1990). Recombinant PCR. In *PCR Protocols: a Guide to Methods and Applications* (Innis, M. A., Gelford, D. H., Sninsky, J. J. & White, T. J., eds), pp. 177–183, Academic Press, Inc., New York.

Horovitz, A. & Fersht, A. R. (1990). Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins. *J. Mol. Biol.* **214**, 613–617.

Hughson, F. M. & Baldwin, R. L. (1989). Use of site-directed mutagenesis to destabilize native apomyoglobin relative to folding intermediates. *Biochemistry*, **28**, 4415–4422.

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 intermediate of apomyoglobin by site-directed mutagenesis. *Biochemistry*, **30**, 4113–4118.

Irace, G., Balestrieri, C., Parlato, G., Servillo, L. & Colonna, G. (1981). Tryptophanyl fluorescence heterogeneity of apomyoglobins. Correlation with the presence of two distinct structural domains. *Biochemistry*, **20**, 792–799.

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.

Kawamura-Konishi, T., Kihara, H. & Suzuki, H. (1988). Reconstitution of myoglobin from apoprotein and heme, monitored by stopped-flow absorption, fluorescence and circular dichroism. *Eur. J. Biochem.* **170**, 589–595.

Kendrew, J. C. (1962). Side-chain interactions in myoglobin. *Brookhaven Symp. Biol.* (*Enzyme Models and Enzyme Structure*) **15**, 216–228.

Kihara, H., Takahashi, E., Yamamura, K. & Tabushi, I. (1980). Kinetic study on the unfolding of apomyoglobin, detected by fluorescence- and circular dichroism-stopped-flow method. *Biochem. Biophys. Res. Commun.* **95**, 1687–1694.

Kirby, E. P. & Steiner, R. F. (1970). The tryptophan microenvironments in apomyoglobin. *J. Biol. Chem.* **254**, 6300–6306.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). *Methods Enzymol.* **154**, 367–382.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures; estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

Lesk, A. M. & Chothia, C. (1980). How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. *J. Mol. Biol.* **136**, 225–270.

McNutt, M., Mullins, L. S., Raushel, F. M. & Pace, C. N. (1990). Contribution of histidine residues to the conformational stability of ribonuclease T1 and mutant Glu58→Ala. *Biochemistry*, **29**, 7572–7576.

Merino, F., Osuna, J., Bolivar, F. & Soberon, X. (1992). A general, PCR-based method for single or combinatorial oligonucleotide-directed mutagenesis on pUC/M13 vectors. *Biotechniques*, **12**, 508–509.

Ptitsyn, O. B. & Rashin, A. A. (1975). A model of myoglobin self-organization. *Biophys. Chem.* **3**, 1–20.

Reynolds, W. F., Peat, I. R., Freedman, M. H. & Lyerla, J. R., Jr (1973). Determination of the tautomeric form of the imidazole ring of L-histidine in basic solution by carbon-13 magnetic resonance spectroscopy. *J. Amer. Chem. Soc.* **95**, 328–331.

Rumen, N. M. & Appella, E. (1962). Molecular association behaviour of apomyoglobin I from the seal (*Phoca vitulina*). *Arch. Biochim. Biophys.* **97**, 128–133.

Sali, D., Bycroft, M. & Fersht, A. R. (1988). Stabilization of protein structure by interaction of α-helix dipole with a charged side chain. *Nature (London)*, **335**, 740–743.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, vols 1–3.

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F. (1981). Determination of nucleotide sequences in DNA. *Science*, **214**, 1205–1210.

Santoro, M. M. & Bolen, D. W. (1992). A test of the linear extrapolation of unfolding free energy changes over an extended denaturation concentration range. *Biochemistry*, **31**, 4901–4907.

Schellman, J. A. (1978). Solvent denaturation. *Biopolymers*, **17**, 1305–1322.

Shen, L. L. & Hermans, J., Jr (1972). Kinetics of conformation change of sperm-whale myoglobin. III. Folding and unfolding of apomyoglobin and the suggested overall mechanism. *Biochemistry*, **11**, 1845–1849.

Shortle, D. & Meeker, A. K. (1986). Mutant forms of staphylococcal nuclease with altered patterns of guanidine hydrochloride and urea denaturation. *Proteins: Struct. Funct. Genet.* **1**, 81–89.

Shikama, K. & Matsuoka, A. (1989). Spectral properties unique to the myoglobins lacking the usual distal histidine residue. *J. Mol. Biol.* **209**, 489–491.

Shin, H.-C., Merutka, G., Waltho. J. P., Tennant, L. L., Dyson. H. J. & Wright, P. E. (1993*a*). Peptide models of protein folding initiation sites. 3. The G-H helical hairpin of myoglobin. *Biochemistry*, **32**, 6356–6364.

Shin, H.-C., Merutka, G., Waltho. J. P., Wright, P. E. & Dyson. H. J. (1993*b*). Peptide models of protein folding initiation sites. 2. The G-H turn region of myoglobin acts as a helix stop signal. *Biochemistry*, **32**, 6348–6355.

Takano, T. (1977). Structure of myoglobin refined at 2·0 Å resolution I. Crystallographic refinement of metmyoglobin from sperm whale. *J. Mol. Biol.* **110**, 537–568.

Teale, F. W. J. (1959). Cleavage of the haem-protein link by acid methylethylketone. *Biochim. Biophys. Acta.* **35**, 543.

Tirado-Rives, J. & Jorgensen, W. L. (1993). Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry*, **32**, 4175–4184.

Tsuji, T., Chrunyk, B. A., Chen, X. & Matthews, C. R. (1993). Mutagenic analysis of the interior packing of an α/β barrel protein. Effects on the stabilities and rates of interconversion of the native and partially folded forms of the α subunit of tryptophan synthase. *Biochemistry*, **32**, 5566–5575.

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 13 August 1993; accepted 4 January 1994)