

# The pK<sub>a</sub> of His-24 in the folding transition state of apomyoglobin

Marc Jamin\*, Bernhard Geierstanger<sup>†</sup>, and Robert L. Baldwin<sup>§</sup>

Department of Biochemistry, Beckman Center, Room B400, Stanford University Medical Center, Stanford, CA 94305-5307

Contributed by Robert L. Baldwin, March 30, 2001

**In native apomyoglobin, His-24 cannot be protonated, although at pH 4 the native protein forms a molten globule folding intermediate in which the histidine residues are readily protonated. The inability to protonate His-24 in the native protein dramatically affects the unfolding/refolding kinetics, as demonstrated by simulations for a simple model. Kinetic data for wild type and for a mutant lacking His-24 are analyzed. The pK<sub>a</sub> values of histidine residues in native apomyoglobin are known from earlier studies, and the average histidine pK<sub>a</sub> in the molten globule is determined from the pH dependence of the equilibrium between the native and molten globule forms. Analysis of the pH-dependent unfolding/refolding kinetics reveals that the average pK<sub>a</sub> of the histidine residues, including His-24, is closely similar in the folding transition state to the value found in the molten globule intermediate. Consequently, protonation of His-24 is not a barrier to refolding of the molten globule to the native protein. Instead, the normal pK<sub>a</sub> of His-24 in the transition state, coupled with its inaccessibility in the native state, promotes fast unfolding at low pH. The analysis of the wild-type results is confirmed and extended by using the wild-type parameters to fit the unfolding kinetics of a mutant lacking His-24.**

His-24 plays a remarkable role in the structure and pH titration behavior of apomyoglobin (apoMb). His-24, which is in the B helix, is completely buried, and its neutral imidazole ring is hydrogen bonded to that of His-119 in the G helix (1–3), thus providing a strong link (4) between the B and G helices. Although His-119 can be protonated at low pH, His-24 remains neutral and cannot be protonated in the native form (N) of apoMb (2). The charged guanidinium group of Arg-118 sits directly above the imidazole ring of His-24 in N (5). Consequently, protonation of His-24 occurs only when N forms a stable molten globule intermediate (I) at pH 4. Only a subdomain of N containing the A, G, and H helices of apoMb is stably folded in I (6).

His-48, His-64, His-113, and His-119 have low pK<sub>a</sub> values in N (2, 7), but only His-24 is entirely resistant to being protonated in N; and protonation of His-24 at pH 4 forces the conversion of N to I. All histidine residues are protonated in the molten globule at and below pH 5, where I is formed (2); thus, the pK<sub>a</sub> values of all histidines in the molten globule are close to normal.

Because His-24 cannot be protonated in N, the question arises of whether His-24 can be protonated in the folding transition state (I<sup>‡</sup>) of the N to I reaction. Barrick *et al.* (4) pointed out that, if His-24 is inaccessible in the folding transition state, this will present a serious kinetic barrier to the refolding of I to N at pH values where I is protonated. To answer this question, we first determine the average pK<sub>a</sub> value of the histidine residues in the molten globule, from the pH-dependent equilibrium between N and I, using data from ref. 8, combined with the pK<sub>a</sub> values of histidine residues in the native state, determined earlier (2, 7). The N ⇌ I equilibrium depends on pH through ΔQ, the difference between the number of protons bound to N vs. I, and the pH dependence of its standard free energy is given by  $-2.303\int\Delta QdpH$ . This integral is calculable from the pK<sub>a</sub>s of the histidines in N and I (9) when unfolding occurs in the pH range in which only histidines are titrating. We use the extensive data

of Barrick and Baldwin (8) on the pH dependence of the equilibrium [N]/[I] ratio.

After determining the average pK<sub>a</sub> value of histidines in the molten globule intermediate, the average pK<sub>a</sub> value of histidines in the folding transition state may be calculated from the plot of log *k*<sub>obs</sub> vs. pH (9, 10), because the pK<sub>a</sub> values in N and I are known. Simulations reveal that, if any histidine has a pK<sub>a</sub> in the folding transition state that differs markedly from the average value, this fact will be apparent in the pH dependence of the unfolding/refolding kinetics.

The double mutant H24V/H119F (4) provides an important control for this analysis of the unfolding/refolding kinetics of wild type (WT). The double mutant is stably folded: the double replacement of His-24 and His-119 with Val and Phe is found commonly in the globin family, and the nonpolar Val-Phe link between the B and G helices makes the double mutant more stable to urea unfolding at neutral pH than WT (4). The single mutants H24V and H119F are both considerably less stable than WT (4). The main cause for the WT native state's becoming unstable at low pH is the protonation of His-24 in I, and, because the double mutant lacks His-24, its N ⇌ I unfolding transition is shifted down near pH 3, where the Asp and Glu residues undergo protonation (2, 4). The unfolding transition of the less stable mutant H36Q, which also is studied here, is shifted in the opposite direction to a higher pH range because protonated His-36 stabilizes WT through side chain interactions (4).

## Experimental Procedures

**Protein Expression and Purification.** A synthetic gene for sperm whale myoglobin was expressed and purified as described (2, 11). The heme was removed by acid-acetone precipitation (12). Protein concentration was determined by absorbance in 6 M GdmCl as described (13) using  $\epsilon_{280\text{ nm}} = 15,200\text{ M}^{-1}\text{cm}^{-1}$  and  $\epsilon_{288\text{ nm}} = 10,800\text{ M}^{-1}\text{cm}^{-1}$ .

**Fluorescence Stopped-Flow Experiments.** Data were collected on an Applied Photophysics (Surrey, U.K.) model DX17 MV instrument, using a 2-mm pathlength cell. Excitation was at 288 nm, and measurements of emission were made by using an optical cutoff filter. Kinetic traces from the stopped-flow experiments were fitted to single, double, and triple exponential time courses with the software provided by Applied Photophysics.

**Modeling the pH Dependence of the Rate Constants.** All simulations were performed with SIGMAPLOT (Jandel, San Rafael, CA). The

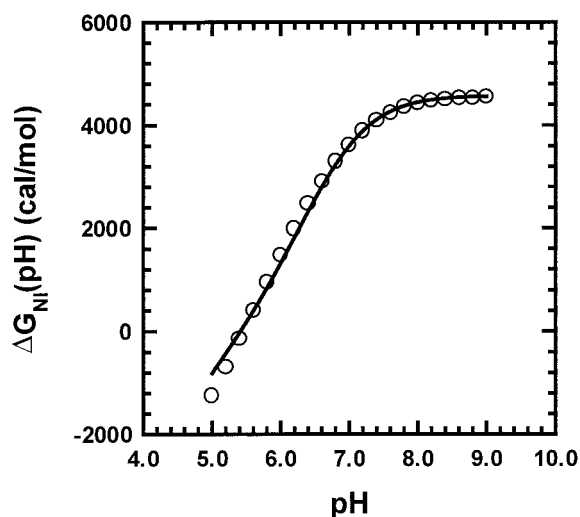
Abbreviations: N, native form; I, folding intermediate; I<sup>‡</sup>, folding transition state; U, acid-unfolded form; WT, wild type; ΔQ, difference in number of protons bound to N and I; pH<sub>m</sub>, pH midpoint of unfolding transition; ΔG<sub>Ni</sub>, difference in standard Gibbs free energy between N and I; apoMb, apomyoglobin.

\*Present address: Laboratoire de Biophysique Moléculaire et Cellulaire, DBMS/BMC CEA-Grenoble, 17 rue des Martyrs, Grenoble 38054, France.

<sup>†</sup>Present address: Genomics Institute of the Novartis Research Foundation, 3115 Merryfield Row, Suite 200, San Diego, CA 92121-1125.

<sup>§</sup>E-mail: bbaldwin@cmgm.stanford.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Fig. 1.** Determination at 0°C of the average pK<sub>a</sub> of histidines in the molten globule (see text for details). Values of ΔG<sub>NI</sub>, the change in Gibbs energy for unfolding from N to I, are taken from ref. 8. The fitted curve is given by  $-2.303 \int \Delta Q \, dpH$ , where ΔQ is the difference in number of protons bound to N vs. I; ΔQ is computed from the pK<sub>a</sub> values in N measured by chemical shift titration (2,7) and from the average pK<sub>a</sub> in I, treated as the sole unknown.

rate constants for folding ( $k_f$ ) and unfolding ( $k_u$ ) are expressed as functions of pH by assuming that the proton binding sites are independent, using the following equations.

$$k_f = k_f^{\text{ref}} \frac{\prod [1 + 10^{-(pK_{ai}^{\ddagger} - pH)}]}{\prod [1 + 10^{-(pK_{ai}^{\text{Ib}} - pH)}]} \quad [1a]$$

$$k_u = k_u^{\text{ref}} \frac{\prod [1 + 10^{-(pK_{ai}^{\ddagger} - pH)}]}{\prod [1 + 10^{-(pK_{ai}^{\text{N}} - pH)}]} \quad [1b]$$

$$K = \frac{k_f}{k_u}, \quad [2]$$

where  $k_f^{\text{ref}}$  and  $k_u^{\text{ref}}$  are the reference rate constants for the  $\text{Ib} \rightarrow \text{N}$  and  $\text{N} \rightarrow \text{Ib}$  reactions, respectively, at pH 9;  $K$  is the folding equilibrium constant; and the pK<sub>ai</sub> values refer to residue  $i$  in the N, Ib, and I<sup>‡</sup> forms.

## Results

**Average pK<sub>a</sub> of Histidines in the Molten Globule.** Fig. 1 shows ΔG<sub>NI</sub>(pH), the difference in standard Gibbs free energy between N and I, calculated from data given by Barrick and Baldwin (8), who used circular dichroism to measure the pH-induced and urea-induced unfolding transitions of apoMb at 0°C from pH 1 to 8 and from 0 to 8 M urea. They fitted the results to a simplified three-state model ( $\text{N} \rightleftharpoons \text{I} \rightleftharpoons \text{U}$ ) using a minimal number of parameters. The urea- and pH-induced unfolding results are fitted closely by the three-state model, and the fitting parameters can be used to give reasonably accurate values of ΔG<sub>NI</sub>(pH) as a function of pH.

In Fig. 1, ΔG<sub>NI</sub>(pH) is plotted against pH, and the fitted line is  $-2.303 \int \Delta Q \, dpH$ , where ΔQ is computed from the known individual pK<sub>a</sub> values of the histidines in N and from the average pK<sub>a</sub> of the histidines in I, treated as the sole unknown. The pK<sub>a</sub> values in N are corrected to 0°C [the temperature used by Barrick and Baldwin (8)] from the 35°C values of Geierstanger *et al.* (2) by using the heat of protonation of free histidine,  $-6.9 \text{ kcal/mol}$  (14), in the van't Hoff relation. The average pK<sub>a</sub> in I is found by least-squares fitting to be 6.8 (0°C). His-36 is omitted from the calculation of  $\int \Delta Q \, dpH$  in Fig. 1 because its protonated form stabilizes N, contrary

to the assumptions of the simple three-state model, and as a result the contribution of His-36 to the experimental  $\text{N} \rightleftharpoons \text{I}$  data is smoothed out in fitting the data to the model of ref. 8. The native pK<sub>a</sub> determinations of histidine residues were made in 10 mM acetate buffer (2) whereas data for the  $\text{N} \rightleftharpoons \text{I}$  unfolding transition (8) were obtained in 2 mM citrate buffer. We show below that, although there is a significant difference in ΔG<sub>NI</sub> in these two buffers, the same pK<sub>a</sub> values apply in both buffers.

**Unfolding and Refolding Kinetics vs. pH.** Although the nature of the rate equation for protein folding is still under discussion (15), there is general agreement that the transition state approximation provides the basis for the rate equation: it specifies that the folding transition state I<sup>‡</sup> is in equilibrium with the denatured protein in refolding and with the native protein in unfolding. There are two forms of the apoMb molten globule, Ia and Ib, which coexist in a pH-dependent equilibrium, and the kinetics of the overall folding reaction indicate that the pathway is  $\text{N} \rightleftharpoons \text{Ib} \rightleftharpoons \text{I}^{\ddagger} \rightarrow \text{N}$ , in which N is formed directly from Ib, not Ia (16); U is the acid-unfolded form. The designation “I” is used here as a collective term for (Ia,Ib) when both forms are present and no further specification is needed.

In refolding to N, I<sup>‡</sup> equilibrates with Ib, and the pH dependence of the folding rate  $k_f$  is attributed to this equilibrium:



$$k_f = K_{\text{I}^{\ddagger}/\text{Ib}} k_{\ddagger\text{N}}. \quad [3b]$$

$K_{\text{I}^{\ddagger}/\text{Ib}}$  is the equilibrium constant for equilibration between Ib and I<sup>‡</sup>, and  $k_{\ddagger\text{N}}$  is the rate constant with which I<sup>‡</sup> breaks down to give N. Likewise, in unfolding to Ib, N equilibrates with I<sup>‡</sup>, and the unfolding rate  $k_u$  is written



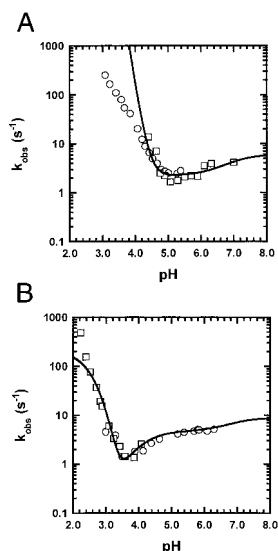
$$k_u = K_{\text{N}/\text{I}^{\ddagger}} k_{\ddagger\text{I}} \quad [4b]$$

In fitting the experimental curve of  $\log k_{\text{obs}}$  vs. pH,  $K_{\text{I}^{\ddagger}/\text{Ib}}$  is pH dependent in refolding experiments, and  $K_{\text{N}/\text{I}^{\ddagger}}$  is pH dependent in unfolding experiments because these equilibrium constants depend on the pK<sub>a</sub> values of residues that titrate in this pH range. For a two-state reaction, the observed rate constant  $k_{\text{obs}}$  is the sum of the unfolding and refolding rate constants.

$$k_{\text{obs}} = k_u + k_f \quad [5]$$

Fig. 2 shows  $k_{\text{obs}}$  plotted against pH at 4.5°C, in citrate buffer, for WT (Fig. 2A) and for the double mutant H24V/H119F (Fig. 2B). The average pK<sub>a</sub> of histidine residues in the folding transition state, determined from the WT curve, is 6.7, almost the same as in the molten globule (6.8, Table 1). Simulations presented below show that, if His-24 had an abnormally low pK<sub>a</sub> in I<sup>‡</sup> as it does in N, this would be readily observable from the shape of the curve of  $\log k_{\text{obs}}$  vs. pH. Exactly the same parameters are used to fit the WT curve and the H24V/H119F curve (Fig. 2B), whereas the pK<sub>a</sub> values of His-24 and His-119 are omitted in the latter. The pK<sub>a</sub> values used to fit these curves (Table 1) are discussed below.

**Results in Acetate Buffer.** Because citrate buffer causes the  $\text{Ia} \rightleftharpoons \text{Ib}$  equilibrium to be more strongly dependent on pH than acetate buffer (ref. 16 and unpublished results), and because we thought this behavior might affect the pH dependence of the unfolding/refolding kinetics, we also determined the unfolding/refolding kinetics in 10 mM acetate buffer. Fig. 3 shows the results for WT (Fig. 3A) and two mutants (H24V/H119F), whose unfolding curve is shifted to lower pH (Fig. 3B), and H36Q (Fig. 3C), whose unfolding curve is shifted to higher pH. These curves are



**Fig. 2.** The plot of  $\log k_{\text{obs}}$  vs. pH for pH-induced unfolding/refolding in a citrate buffer (2 mM citrate/30 mM NaCl) at 4.5°C for WT (A) and the double mutant H24V/H119F (B). Both mutant and WT curves are fitted by using the same  $\text{pK}_a$  values (Table 1), with reference values  $k_u = 0.001$  and  $k_f = 6$  (A) or with  $k_u = 2 \times 10^{-5}$  and  $k_f = 9$  (B). The rate constants are given in units of  $\text{s}^{-1}$  and the reference values refer to pH 9.

fitted by the same  $\text{pK}_a$  values determined in citrate (Table 1), whereas the altered stability of I in acetate (Table 2) is taken into account by adjusting the values of  $k_u$  and  $k_f$ .

In analyzing the kinetics of a folding reaction, it is necessary to examine the kinetic amplitudes as well as the rates. The kinetic amplitudes in acetate buffer are shown in Fig. 4A for WT and in Fig. 4B for H24V/H119F. The unfolding or refolding reactions start at pH values where either N (unfolding) or I (refolding) is present in the initial conditions, and the pH-jump typically ends at pH values inside the unfolding transition zone, so that the final fluorescence value follows the equilibrium unfolding transition if the reaction is

**Table 1.  $\text{pK}_a$  values used to fit the pH-dependent unfolding/refolding kinetics**

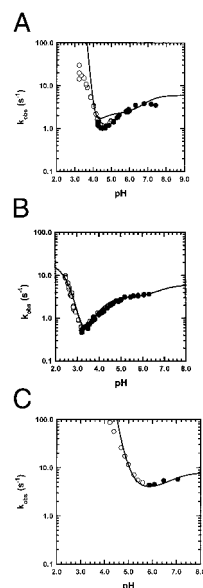
Residue	$\text{pK}_a$ , N*	$\text{pK}_a$ , I <sup>††</sup>	$\text{pK}_a$ , I <sup>§</sup>	No.
His-24	>3.5	6.7	6.8	1
His-36	8.5	6.7	6.8	1
His-48	6.0	6.7	6.8	1
His-64	5.5	6.7	6.8	1
His-113	6.5	6.7	6.8	1
His-119	6.5	6.7	6.8	1
Acidic I	3.4	4.0	4.0	9
Acidic II	3.4	3.5	3.6	13

All  $\text{pK}_a$  values refer to 4.5°C, the temperature of the experiments reported here. The estimated uncertainty of the histidine  $\text{pK}_a$  values is  $\pm 0.1$ . See text for the procedure used to fix  $\text{pK}_a$  values of the acidic residues (Asp and Glu): group I corresponds to the unfolded CDEF subdomain of the apoMb molten globule, and group II corresponds to the structured A[B]GH subdomain of the molten globule.

\*The  $\text{pK}_a$  values of histidine residues in native apoMb are taken from table 2 of ref. 2, determined at 35°C and corrected to 4.5°C (see text). No correction has been made for the possible influence of D<sub>2</sub>O on the histidine  $\text{pK}_a$  values determined by chemical shift titration.

†The average  $\text{pK}_a$  value of histidine residues in the folding transition state, determined by fitting the plot of  $\log k_{\text{obs}}$  versus pH for wild type in citrate buffer (Fig. 2A), when the  $\text{pK}_a$  values in N and in I<sub>b</sub> are known.

§The average  $\text{pK}_a$  of the histidine residues in the molten globule, determined as shown in Fig. 1.



**Fig. 3.** The plot of  $\log k_{\text{obs}}$  vs. pH in 10 mM acetate buffer at 4.5°C for WT (A), H24V/H119F (B), and H36Q (C). All three curves are fitted with same  $\text{pK}_a$  values (Table 1), which are also used to fit the data in citrate buffer (Fig. 2), with reference values  $k_u = 2 \times 10^{-5}$ ,  $k_f = 6$  (A),  $k_u = 2 \times 10^{-6}$ ,  $k_f = 6$  (B),  $k_u = 5 \times 10^{-4}$ , and  $k_f = 8$  (C). The rate constants are given in  $\text{s}^{-1}$ , and the reference values refer to pH 9.

two-state. The fluorescence attributed to the two Trp residues in apoMb increases when N unfolds to I because of specific fluorescence quenching in N. The open and filled circles show the initial and final fluorescence values in refolding, whereas the open and filled squares show the initial and final values in unfolding. The final fluorescence values follow the equilibrium unfolding curve of N to I at pH values where only N and I are present, but, at lower pH values, U is also formed and the fluorescence again decreases. U has a lower fluorescence than I because the two Trp residues of apoMb are partially sequestered from solvent in I but are solvent exposed in U (16). The equilibrium pH unfolding curve obtained by monitoring fluorescence is subject to some uncertainty because of the difficulty of determining the correct unfolded and native base lines.

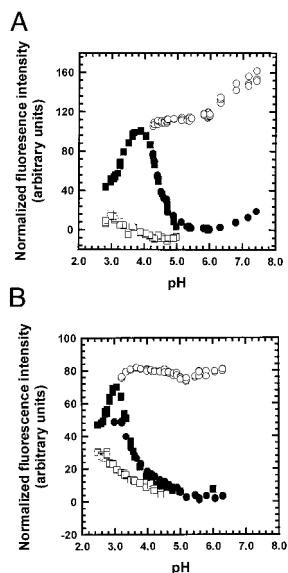
**Quality of the Data Fitting.** By using a single set of  $\text{pK}_a$  values to fit the pH-dependent unfolding/refolding kinetics not only of WT but also of the two mutants that unfold in different pH ranges (Fig. 3 B and C), we have a further test of the  $\text{pK}_a$  values. Likewise, using the same  $\text{pK}_a$  values to fit the kinetic data taken

**Table 2. Properties of the pH-induced unfolding reactions in citrate vs. acetate buffer**

Protein	Citrate		Acetate	
	$\text{pH}_m$	$\Delta G_{\text{NI}}$	$\text{pH}_m$	$\Delta G_{\text{NI}}$
WT	4.7	4.8	4.4	7.0
H24V/H119F	3.5*	7.2	3.2†	8.2
H36Q	n.d.	n.d.	5.3	5.3

Determined at 4.5°C in either 10 mM acetate or 2 mM citrate, 30 mM NaCl buffer.  $\text{pH}_m$  is the pH midpoint of the unfolding transition monitored by the final fluorescence values obtained in unfolding/refolding experiments, as in Fig. 4 A and B.  $\Delta G_{\text{NI}}$  is the Gibbs energy difference between the native and molten globule forms at pH 9, in kcal/mol, computed from the values of  $k_u$  and  $k_f$  used to fit the curve of  $\log k_{\text{obs}}$  vs. pH.

\*A lower value than the one given here is obtained if the baseline found for the WT protein is used.

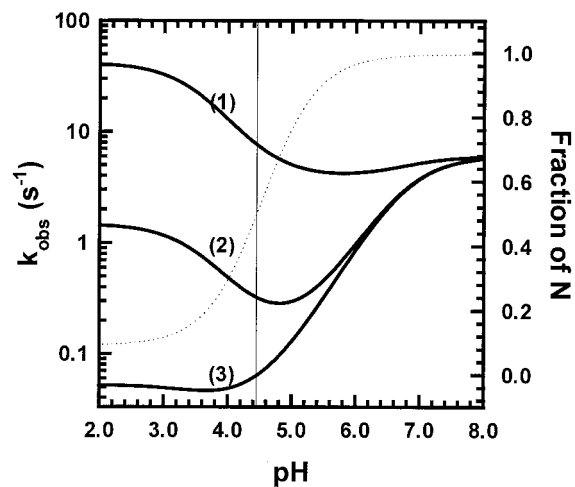


**Fig. 4.** The kinetic amplitudes of the unfolding and refolding experiments shown in Fig. 3 *A* and *B*. Open symbols are initial values, filled symbols are final values. Circles are refolding starting from pH 4.2, and squares are unfolding starting from pH 6.0. The fluorescence data are normalized so that the final fluorescence is 0 at pH 6.0 and 100 at pH 4.2. The fitted curves are equilibrium pH unfolding curves, which yield two of the  $pH_m$  values shown in Table 2

in both acetate and citrate buffers provides an additional test of the  $pK_a$  values, because the location of the pH unfolding curve shifts significantly between acetate and citrate for both WT and H24V/H119F (Table 2). By insisting that a single set of  $pK_a$  values should fit all of the data, and by determining, one step at a time, the average  $pK_a$  values of the histidine residues in the I $\ddagger$  intermediate and in the folding transition state, we have a severe test of the assumptions made in this analysis as well as of the  $pK_a$  values. The fact that the fitted curve appears to be better for the double mutant (Figs. 2*B* and 3*B*) than for WT (Figs. 2*A* and 3*A*) probably reflects the fact that only one  $pK_a$  (the average  $pK_a$  of I $\ddagger$ ) is an adjustable parameter in fitting the WT data, whereas the acidic  $pK_a$  values of both subdomains (Table 1) are adjusted in fitting the H24V/H119F data. The fitted curves for WT (Figs. 2*A* and 3*A*) diverge from the data points at low pH. The divergence can be eliminated by abolishing the carboxylate  $pK_a$  differences between different forms of the protein (I, I $\ddagger$ , and N), but then it becomes impossible to fit the pH-dependent unfolding/refolding kinetics of H24V/H119F. The divergence occurs at pH values where unfolding ceases to be a two-state process, as the acid-unfolded form U becomes populated (compare Figs. 3*A* and 4*A*), and we tentatively assign the divergence to failure of the two-state assumption in this pH range. A second assumption that likely affects the quality of the fitted data is that the  $pK_a$  values are independent of the extent of titration of neighboring groups.

#### Simulation of the Effect of His-24 on the Unfolding/Refolding Kinetics.

The fact that His-24 cannot be protonated in N gives this residue a unique role in the unfolding/refolding kinetics of WT, as shown by simulations for a simple model with only one ionizing residue. The  $pK_a$  of this residue in N is so low that its protonation cannot be observed. The rate curves in Fig. 5 are simulated for three different  $pK_a$  values in I $\ddagger$ : either 6.6 (curve 1), 5.15 (curve 2), or 3.7 (curve 3). The  $pK_a$  of the single residue is set at 3.5 in N and at 6.8 in I. The assigned  $pK_a$  of 3.5 in N does not allow this residue to be observably protonated in N because I is formed before N is protonated. The equilibrium transition curve for pH unfolding (dotted line in Fig. 5) is given by the ratio  $k_f/k_u$ ,



**Fig. 5.** Simulation of the values of  $k_{obs}$  vs. pH for a simple model (see text) in which only one ionizing group has  $pK_a$  values that differ in I, I $\ddagger$ , and N. The  $pK_a$  value in I $\ddagger$  differs in the three cases shown: case 1, 6.6 (close to I); case 2, 5.15 (intermediate); case 3, 3.7 (close to N). The  $pK_a$  value in N is 3.5 and in I is 6.8 in all three cases. The fraction unfolded is shown on the right y axis, and the dotted line is the equilibrium unfolding curve (unfolding does not go to completion at low pH) calculated from the values of  $k_u$  and  $k_f$  vs. pH (see *Experimental Procedures*). Their values at pH 9 are  $k_u = 0.001 s^{-1}$  and  $k_f = 6 s^{-1}$ .

containing the rate constants for folding and unfolding, and  $pH_m$ , the pH midpoint of the equilibrium unfolding curve, is near 4.4, close to the experimental  $pH_m$  for WT in acetate. The unfolding transition does not go to completion at low pH because the assigned  $pK_a$  difference between N and I is not large enough to accomplish this when only a single residue is involved. The part of each simulated rate curve (Fig. 5) that lies below  $pH_m$  contributes to the unfolding rate whereas the part that lies above  $pH_m$  contributes to the refolding rate. The simulated curves show that the presence of a single residue whose  $pK_a$  differs substantially between N and I will have a large effect on the pH dependence of the unfolding/refolding kinetics.

The main qualitative features of the simulated results can be understood by noting two properties of the folding rate equation. First, at pH values below  $pH_m$ , the unfolding rate typically dominates the overall rate (the sum of the unfolding and refolding rates), whereas above  $pH_m$  the reverse is true. Second, the  $pK_a$  difference between N and I $\ddagger$  controls the unfolding rate whereas the refolding rate is controlled by the  $pK_a$  difference between I and I $\ddagger$ . Thus, the refolding rate is nearly pH independent in case 1 ( $pK_a$  of I $\ddagger$  is close to I), whereas the unfolding rate is nearly pH independent in case 3 ( $pK_a$  of I $\ddagger$  is close to N). In case 2, ( $pK_a$  of I $\ddagger$  intermediate between N and I) both the unfolding and refolding rates are strongly pH dependent. In experiments on pH-induced unfolding, the minimum value of the observed rate normally falls near the midpoint of the transition curve, as in denaturant-induced unfolding. The rate curve becomes independent of pH when the difference in number of protons bound becomes pH independent: for N versus I $\ddagger$  if unfolding is being measured, and for I versus I $\ddagger$  if refolding is being measured.

#### Discussion

##### Histidine $pK_a$ Values Are Close to Model Compound Values in both I and I $\ddagger$

The average  $pK_a$  of the histidine residues in the transition state (6.7) is close to the average  $pK_a$  in the molten globule (6.8; Table 1), and both  $pK_a$  values are determined by procedures in which only a single adjustable parameter is determined. The average histidine  $pK_a$  value determined for the molten globule is reasonably close to values measured in small peptides [6.8–6.9 at 25°C (17) or 7.2–7.3

at 4.5°C], considering that histidine pK<sub>a</sub> values are affected significantly by neighboring charged residues (17, 18). In native apoMb, the range of pK<sub>a</sub> values covers 3 pH units even when His-24 is excluded (see Table 1). Presumably the histidine residues in the molten globule are well exposed to solvent and do not make strong interactions with neighboring side chains, unlike histidine residues in the native protein. The effect of net charge on the pK<sub>a</sub> values of the molten globule intermediate should be small at the pK<sub>a</sub> of the histidine residues in I because the net proton charge is small: there are 15 basic groups, including histidines, and 13 acidic groups in the A[B]GH subdomain of apoMb. On the other hand, the pK<sub>a</sub> values of His residues in holomyoglobin are affected by neighboring charged residues (17, 18), and they increase measurably with NaCl concentration (18).

**Nature of the pH-Dependent Unfolding Kinetics of H24V/H119F.** The double mutant H24V/H119F lacks His-24, and so its curve of log *k*<sub>obs</sub> vs. pH might be expected to have a different shape from that of WT. Surprisingly, the two curves are similar in shape (Figs. 2 and 3). The Asp residues of native apoMb were studied by chemical shift titration of <sup>13</sup>C-labeled protein; 5 of 7 Asp residues could be observed and were found to titrate with normal pK<sub>a</sub> values (2). No data for the Glu residues are available. The double mutant kinetic data can be fitted by assigning average pK<sub>a</sub> values to the 9 Glu and Asp residues in the unfolded CDEF subdomain that are 0.6 pH units lower in N than in I<sup>‡</sup> (Table 1). The effect on the pH-dependent unfolding rate is similar to that produced by the large (>3 pH) displacement of the pK<sub>a</sub> of a single residue, His-24, in WT. The smaller increase in the refolding rate of H24V/H119F above pH<sub>m</sub> can be fitted by assigning an average pK<sub>a</sub> in I<sup>‡</sup> to the 13 acidic groups in the A[B]GH subdomain that is only 0.1 pH units less than the average value in I.

The pK<sub>a</sub> values assigned to the Asp and Glu residues (Table 1) are chosen on the basis of an earlier study of the acid-induced unfolding of the apoMb molten globule (19) that found that the Linderstrøm-Lang smeared-charge model describes reasonably well the effect of net positive charge on the stability of the molten globule, with *w* (the sole adjustable parameter of the model) = 0.05 and with an average value of 4.0 assigned to the intrinsic pK<sub>a</sub> values of Asp and Glu. In that study, a set of single mutants, each with one basic residue mutated to alanine, consistently increased the stability of the molten globule if the positively charged group was inside the A[B]GH subdomain. Some additional mutants were studied later, and consistent results were found (20). The pK<sub>a</sub> value of 3.6 assigned to 13 acidic residues (6 Asp, 7 Glu) in the A[B]GH subdomain is the value found from the Linderstrøm-Lang equation with *w* = 0.05. It agrees with the value (pK<sub>a</sub> = 3.54) found by fitting data for the I ⇌ U unfolding transition in figure 2 of ref. 21, by using equation 6 of ref. 2. The CDEF subdomain is treated as being unfolded, and its 9 acidic residues (1 Asp, 7 Glu and 1 C-terminal carboxylate included as being unfolded) are assigned a pK<sub>a</sub> = 4.0.

Then the values assigned to the average pK<sub>a</sub> values in I<sup>‡</sup> and N are obtained by fitting the experimental results.

The shift in pH<sub>m</sub> between acetate and citrate (Table 2) is observed both with WT, whose unfolding occurs in a pH range where His-residues are titrating, and with H24V/H119F, whose unfolding occurs in a pH range where Asp and Glu residues are titrating. The similar behavior found in these two pH ranges indicates that the effect of anion binding is not restricted to one class of ionizing groups, and is consistent with the effect being one of net charge on molten globule stability. Preferential binding of the citrate anion, which would stabilize the molten globule by reducing its net positive charge (22), may account for the smaller value of ΔG<sub>Ni</sub> found in citrate versus acetate. Goto and Nishikori (23) carried out partial acetylation of lysine amino groups in cytochrome *c* and separated the partially reacted species by net charge using isoelectric focusing; they found that molten globule stability at pH 1.8 varies with net positive charge as expected if anion binding stabilizes molten globules.

**Comparison with Related Work.** The effect of protonating one residue (Asp 93) on the pH dependence of the unfolding rate of barnase was studied by Oliveberg and Fersht (10). They found a large rise in unfolding rate below pH<sub>m</sub> and also a comparably large rise in refolding rate above pH<sub>m</sub>, implying that the pK<sub>a</sub> of Asp 93 in I<sup>‡</sup> is roughly half way between its values in the native and denatured states. Our results are quite different, probably because there are important structural differences between barnase and apoMb. Charged Asp-93 in barnase forms part of a buried salt link with Arg-69, and its pK<sub>a</sub> is strongly shifted in N, down to 0.7 in 50 mM KCl. The fact that the pK<sub>a</sub> of Asp-93 in the folding transition state is intermediate between its values in the native and denatured states indicates that the ion pair interaction is formed but not fully developed in the transition state. In contrast to charged Asp 93 in barnase, neutral His-24 in apoMb forms a hydrogen-bonded interaction with His-119.

**Concluding Comment.** Residues with depressed pK<sub>a</sub> values in the acid range are found commonly in studies of the pH dependence of stability of native proteins (24). They are thought to be a main contributor to the behavior, commonly observed in globular proteins, in which thermal stability drops sharply at acid pH (25). Our results suggest that the type of pH dependence found here for the N ⇌ I<sub>b</sub> reaction of apoMb, with a sharp rise in unfolding rate below pH<sub>m</sub> and a smaller rise in refolding rate above pH<sub>m</sub>, is likely to be found commonly in the kinetics of pH-dependent unfolding reactions of small proteins at acid pH, although the barnase results of Oliveberg and Fersht (10) show that this situation will not always occur.

This work was supported in part by research grants from the National Institutes of Health (GM 19988 to R.L.B.) and from the Deutsche Forschungsgemeinschaft (Ge 868/1-1 to B.H.G.).

- Bashford, D., Case, D. A., Dalvit, C., Tennant, L. & Wright, P. E. (1993) *Biochemistry* **32**, 8045–8056.
- Geierstanger, B., Jamin, M., Volkman, B. F. & Baldwin, R. L. (1998) *Biochemistry* **37**, 4254–4265.
- Hennig, M. & Geierstanger, B. (1999) *J. Am. Chem. Soc.* **121**, 5123–5126.
- Barrick, D., Hughson, F. M. & Baldwin, R. L. (1994) *J. Mol. Biol.* **237**, 588–601.
- Takano, T. (1977) *J. Mol. Biol.* **110**, 537–568.
- Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990) *Science* **249**, 1544–1548.
- Cocco, M. J. & Lecomte, J. T. (1994) *Protein Sci.* **3**, 267–281.
- Barrick, D. & Baldwin, R. L. (1993) *Biochemistry* **32**, 3790–3796.
- Tanford, C. (1970) *Adv. Protein Chem.* **24**, 1–95.
- Oliveberg, M. & Fersht, A. R. (1996) *Biochemistry* **35**, 6795–6805.
- Loh, S. N., Kay, M. S. & Baldwin, R. L. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5446–5450.
- Rossi-Fanelli, A., Antonini, E. & Caputo, A. (1958) *Biochim. Biophys. Acta* **30**, 608–615.
- Edelhoch, H. (1967) *Biochemistry* **6**, 1948–1954.
- Edsall, J. T. & Wyman, J. (1958) *Biophysical Chemistry* (Academic, New York).
- Eaton, W. A., Muñoz, V., Hagen, S. J., Jas, G. S., Henry, E. R. & Hofrichter, J. (2000) *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359.
- Jamin, M. & Baldwin, R. L. (1998) *J. Mol. Biol.* **276**, 491–504.
- Takahashi, T., Nakamura, H. & Wada, A. (1992) *Biopolymers* **32**, 897–909.
- Kao, Y.-H., Fitch, C. A., Bhattacharya, S., Sarkisian, C. J., Lecomte, J. T. & Garcia-Moreno, B. (2000) *Biophys. J.* **79**, 1637–1654.
- Kay, M. S. & Baldwin, R. L. (1998) *Biochemistry* **37**, 7859–7868.
- Ramos, C. H. I., Kay, M. S. & Baldwin, R. L. (1999) *Biochemistry* **38**, 9783–9790.
- Jamin, M., Antalík, M., Loh, S. N., Bolen, D. W. & Baldwin, R. L. (2000) *Protein Sci.* **9**, 1340–1346.
- Goto, Y., Takahashi, N. & Fink, A. L. (1990) *Biochemistry* **29**, 3480–3488.
- Goto, Y. & Nishikori, S. (1991) *J. Mol. Biol.* **222**, 679–686.
- Anderson, D. E., Lu, J., McIntosh, L. & Dahlquist, F. W. (1993) in *NMR of Proteins: Topics in Molecular and Structural Biology*, eds. Clore, G. M. & Gronenborn, A. M. (CRC, Boca Raton), pp. 258–304.
- Privalov, P. L. (1979) *Adv. Protein Chem.* **33**, 167–241.