J. Mol. Biol. (1998) **276**, 491–504

JMB

Two Forms of the pH 4 Folding Intermediate of Apomyoglobin

Marc Jamin and Robert L. Baldwin*

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

**Corresponding author*

The pH 4 folding intermediate of apomyoglobin exists in two forms (Ia, Ib) at equilibrium. Their ratio depends on pH, urea concentration and the presence or absence of a stabilizing anion (citrate, sulfate), and it does not depend on protein concentration. The Ia and Ib species are separated by a kinetic barrier and their interconversion can be monitored by tryptophan fluorescence in stopped-flow experiments. At pH 4.2, Ib is converted to Ia at low urea concentrations and urea unfolding gives the unfolding transition of Ia.

During the refolding of native (N) apomyoglobin at pH 6, starting from the acid unfolded species (U), both Ia and Ib appear as transient intermediates and both Ia and Ib appear as transient intermediates in the acid-induced unfolding of N. The results are consistent with a linear folding and unfolding pathway: U ⇌ Ia ⇌ Ib ⇌ N. Apomyoglobin provides the opportunity to investigate at equilibrium the structures and properties of two different kinetic folding intermediates.

A non-obligatory dimeric species of the pH 4 intermediate is formed slowly and contributes to the refolding kinetics at concentrations above 5 μM. The dimer dissociates slowly and during refolding at pH 6 it forms N in a later time range than does the monomer.

Keywords: folding intermediate; apomyoglobin; folding kinetics; stopped-flow

Introduction

The pH 4 folding intermediate (I) of sperm whale apomyoglobin (apoMb) (Griko *et al.*, 1988) has properties of particular interest for investigating the mechanism of protein folding. Its helical structure, which has been characterized by NMR–hydrogen exchange (Hughson *et al.*, 1990), suggests that I may be a subdomain of apomyoglobin because I contains the A, G and H helices, which form a compact subdomain in Mb. The A and H helices come from opposite ends of the polypeptide chain, which suggests that the interaction between helices is specific. Nevertheless, I has properties characteristic of molten globule folding intermediates and these intermediates are commonly assumed not to have specific tertiary interactions. Thus, determining whether I does or does not have specific tertiary structure is a main focus of current work on the pH 4 intermediate. I occurs in the refolding kinetics of N at pH 6.0 as a rapidly formed transient intermediate (Jennings & Wright, 1993). Thus, characterization of I at equilibrium at pH 4 provides information about the first known kinetic intermediate in the refolding of N at pH 6. The structure of I changes with conditions: the B helix is recruited into I when the stabilizing anion trichloroacetate (20 mM) is added (Loh *et al.*, 1995). Nevertheless, I unfolds and refolds cooperatively in the urea-induced unfolding transition at pH 4.2 (Kay & Baldwin, 1996; Jamin & Baldwin, 1996).

We report here that two forms of I coexist at pH 4 in an equilibrium that depends on pH, urea concentration and the presence or absence of a stabilizing anion. This finding has basic implications for understanding the folding pathway of apoMb.

Abbreviations used: apoMb, apomyoglobin; GdmCl, guanidinium chloride; Mb, myoglobin; TCA, trichloroacetate; Trp, tryptophan.

Results

Evidence for two forms of I

The fluorescence emission spectrum of I has a substantially greater intensity than that of U, the

0022–2836/98/070491–14 $25.00/0/mb971543

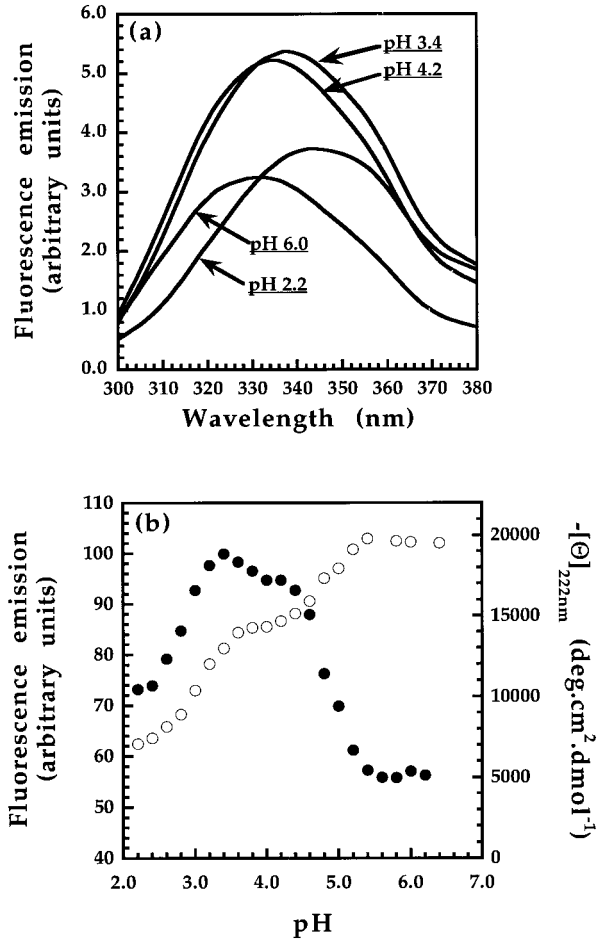

Figure 1. Acid-induced unfolding transition monitored by fluorescence and circular dichroism. Conditions: 2 mM Na citrate/citric acid, 30 mM NaCl, 4.5°C, protein concentration 1 μM. (a) Fluorescence emission spectra of the unfolded form (U, pH 2.2), the native form (N, pH 6.0), and the intermediate at pH 3.4 and 4.2. (b) The transition monitored by CD (○) and fluorescence (●). The integral of the fluorescence intensity between 320 and 380 nm is normalized so that the value at pH 3.4 is 100.

acid-unfolded form, and the wavelength of maximum emission of I is about half way between those of U and N (Figure 1(a)). These properties suggest that the two Trp residues of Mb are partly buried in I, away from solvent; both residues W7 and W14 are in the A helix.

The shape of the fluorescence emission spectrum of I changes with pH between pH 3.4 and 4.2 (Figure 1(a)). The spectrum is slightly red-shifted at pH 3.4 and the maximum intensity is somewhat greater at pH 3.4 than at 4.2. The shape of the pH unfolding curve (monitored by the integral of the fluorescence intensity from 320 to 380 nm) suggests that two forms of I are present, although the unfolding curve monitored by CD reveals only one intermediate (Figure 1(b)). The change in intensity of the spectrum at pH 3.4 cannot be explained as a mixture of I and U because the spectrum of U has much lower intensity than that of I and the admixture of U would result in lower intensity. Also, the change in the spectrum between pH 4.2 and 3.4 cannot be explained as an intrinsic property of the tyrosine or tryptophan residues in I because no change is observed in this pH range with a 2:3 molar ratio of N-acetyl-tryptophanamide and N-acetyl-tyrosinamide in the same concentration range used to study I (data not shown).

Clear evidence for two forms of I in equilibrium with each other is provided by stopped-flow kinetic experiments. The change in fluorescence between pH 3.4 and 4.2 occurs with measurable kinetics (Figure 2(a)) and the data can be fitted to a single exponential which accounts for the observed kinetic amplitude. Matched pH-jump experiments in the forward and reverse directions have equal and opposite amplitudes. The variation of the reaction rate with pH is shown in Figure 2(b) and the pH profile of the change in fluorescence intensity, obtained from the kinetic amplitudes, is shown in Figure 2(c). Thus, the Ia ⇌ Ib reaction is a kinetically reversible two-state reaction.

Urea dependence of the Ia ⇌ Ib reaction

We reported earlier that the kinetics of the urea-induced unfolding–refolding transition at pH 4.2 are measurable when monitored by tryptophan fluorescence, in the time range below 10 ms (Jamin & Baldwin, 1996). Thus, the next question is whether the Ia ⇌ Ib reaction can be separated from the Ia ⇌ U reaction. The equilibrium transition curve for urea-induced unfolding is shown in Figure 3(a) monitored by fluorescence and also by CD. There is a curious rise in fluorescence intensity from 0 to 1 M urea before the intensity falls between 1 and 3 M as unfolding occurs. The rise between 0 and 1 M has been noted before (Kay & Baldwin, 1996; Jamin & Baldwin, 1996), but not explained. Figure 3(b) shows the apparent rate constant for the fluorescence change over the entire range from 0 to 4 M urea; the data above 1 M are taken from Jamin & Baldwin (1996). They observed that the refolding reaction becomes too fast to measure when the urea concentration falls below 1.4 M and they did not investigate the region from 0 to 1 M urea. Figure 3(b) shows that there is in fact a measurable reaction between 0 and 1 M urea; its rate increases as the urea concentration increases, which is the behavior observed in unfolding reactions. The reaction rate cannot be extrapolated accurately to 0 M urea because the amplitude falls rapidly below 0.5 M, but the rate is in the same range as that of the Ia ⇌ Ib reaction measured in pH-jump experiments at pH 4.2 (Figure 2(c)). If Ia is specified to be the form with the higher fluorescence and the red-shifted spectrum, then the reaction observed as the urea concentration increases from 0 to 1 M urea is Ib → Ia.

A test for the presence of Ib in specified initial conditions is to make a stopped-flow experiment

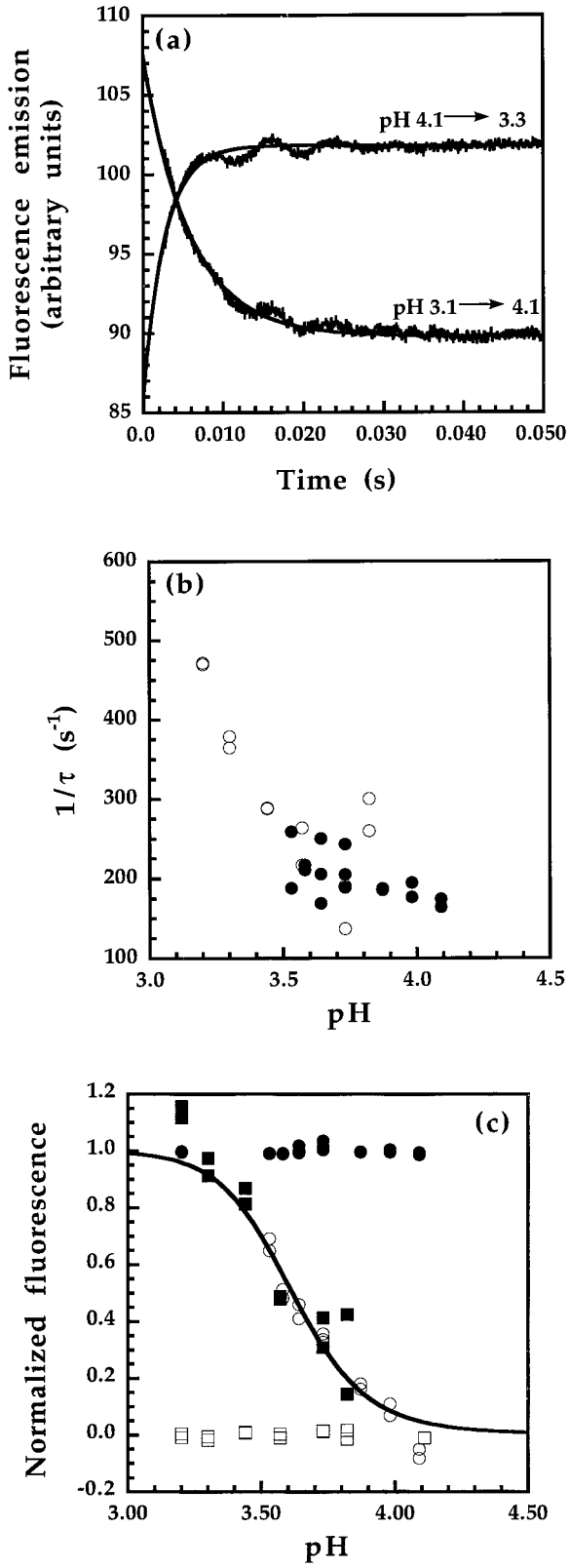

Figure 2. Kinetics of the pH-induced Ia ⇌ Ib interconversion. The starting pH is either 3.1 or 4.1. Representative time traces for jumps from pH 3.1 to pH 4.1 and from pH 4.1 to pH 3.3 are shown. Conditions: 2 mM Na citrate/citric acid, 30 mM NaCl, 4.5°C, final protein concentration is 10 μM after 1:1 dilution. (a) Kinetic traces fitted to a single exponential equation. At higher initial concentrations, a small fluorescence decrease occurs around one second (see text). Otherwise, protein dilution experiments at the same initial and final pH give horizontal traces that superimpose on the level portions of these curves. (b) The reaction rate ($1/\tau$) *versus* final pH: (○) unfolding, (●) refolding. (c) Initial and final fluorescence values: the (○) and (■) show the final kinetic values of fluorescence emission in refolding and unfolding respectively. The (●) and (□) show the initial fluorescence values in the refolding and unfolding kinetics obtained by fitting the kinetic traces to a single exponential and extrapolating the curve to time zero. Linear baselines have been calculated for the initial fluorescence values and data have been normalized after subtraction of the baselines.

with final conditions 0.5 M urea, pH 4.2, and to observe whether or not the Ib → Ia reaction occurs. This test shows that Ib is present in citrate (2 mM) or sulfate (20 mM) at pH 4.2. A related test, now for the formation of Ib in specified final conditions, is to check for the presence of the Ia → Ib reaction in these conditions. This test shows that a small amount of Ib is formed during refolding to N at pH 6.0 in 10 mM Na acetate, after a pH jump from 4.2 to 6.0 (see below).

Ia and Ib occur in the refolding and unfolding kinetics of N

The refolding kinetics of N at pH 6.0, starting either from U (pH 2.2, Figure 4(a)) or from mixtures of Ia and Ib at pH 3.4 (Figure 4(b)) or pH 4.2 (Figure 4(c)), show three measurable kinetic phases plus a burst phase. The three measured kinetic phases are nearly the same for all three starting pH values (Table 1). Starting from U at pH 2.2 gives an additional burst phase which yields a species (Ia) with the same fluorescence intensity as does starting from pH 3.4. The refolding kinetics of N have been measured at varying final pH values, starting from the fixed initial pH 4.2 (data not shown). By continuity, the Ia ⇌ Ib reaction observed in pH jump experiments (Figure 2) is also the first observed kinetic phase in the refolding of N, while the U ⇌ Ia reaction occurs in the burst phase. Consequently, the Ia, Ib folding intermediates observed at equilibrium between pH 3.1 and 4.2 are also observed as kinetic intermediates in the formation of N at pH 6.0, starting from U at pH 2.2.

To find out which kinetic phase (or phases) yields N, a specific assay for N is used, namely its unfolding rate (Schmid, 1983; Kiefhaber, 1995). In an interrupted refolding experiment, the kinetics of unfolding are measured after a pH jump from 6.0 to 4.2 at various times of refolding. The native protein unfolds more slowly than the refolding intermediates. The refolding results, assayed by unfolding rates, are shown in Figure 5. Native apoMb is formed in each of the last two kinetic phases and a refolding intermediate (Ib) is formed in the first phase. The minor final phase, in which

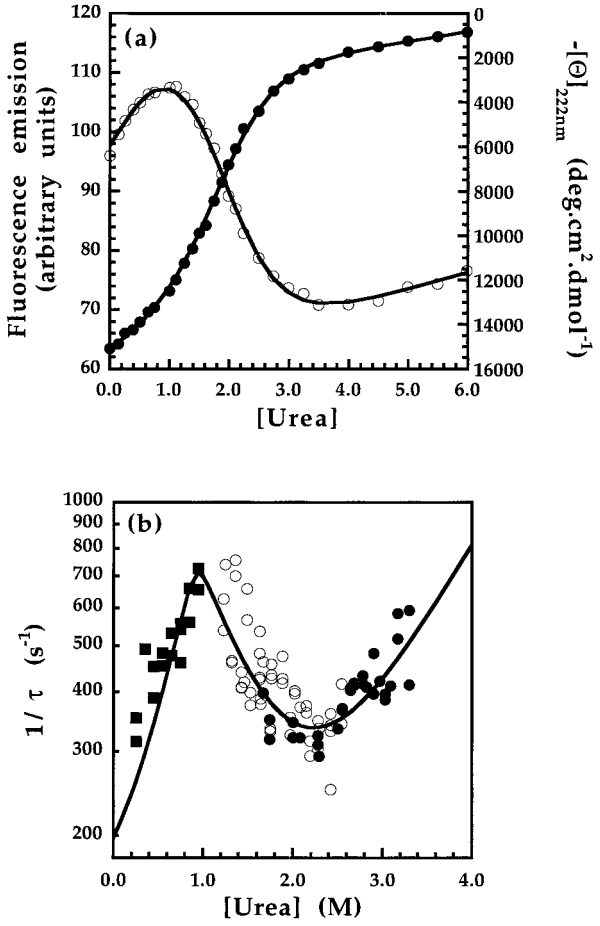

Figure 3. Urea-induced unfolding transition at pH 4.2. Conditions: 2 mM Na citrate/citric acid, 30 mM NaCl, pH 4.2, 4.5°C, final protein concentration 10 μM. (a) Equilibrium transition curves for urea-induced unfolding monitored by far-UV CD (●) and by Trp fluorescence (○). The fluorescence values are normalized so that the value at pH 4.2, 0 M urea is 96. (b) The apparent rate constant ($1/\tau$) *versus* [urea] for refolding from 4 M urea (○) and for unfolding from 0.7 M urea (●) or 0 M urea (■). The data below 1 M urea (■) were obtained in unfolding experiments starting from 0 M urea . The lines are drawn from a combined fit (see methods) of the equilibrium and kinetic data, using the linear mechanism U ⇌ Ia ⇌ Ib and the same parameters in both (a) and (b).

N is formed, represents the dissociation of an oligomeric intermediate and formation of native monomer, according to experiments given below.

The corresponding interrupted unfolding experiment, initiated by a pH jump from 6.0 to 4.2, is stopped at various times of unfolding in order to examine the refolding kinetics (pH 4.2 → 6.0) of the unfolding intermediates. The results are shown in Figure 6(a). The refolding amplitudes, measured at various unfolding times, yield the kinetics for accumulation of both intermediates Ia (filled circles, A_1) and Ib (open circles, A_2). Ia is measured as the amplitude of phase 1 in the refolding kinetics (see legend to Figure 5) and Ib is given by the amplitude of phase 2. The kinetic phase in unfolding with the least amplitude (triangles, A_3, phase 3 in refolding kinetics) yields a dimeric or oligomeric form of I. The ratio of the amplitudes of the first two phases (A_1/A_2) increases in the first 100 ms of unfolding (Figure 6(b)) as expected if Ib must be formed from N before Ia can be formed from Ib. The line shows the simulated behavior of the linear mechanism N ⇌ Ib ⇌ Ia. The mechanism in which Ia and Ib are intermediates on parallel pathways can be ruled out, because the amplitude ratio A_1/A_2 does not change with time for this mechanism. The refolding rates of the three species present at various times of unfolding do not vary with unfolding time (data not shown).

An additional dimeric or oligomeric intermediate

A standard test for the presence of dimeric or oligomeric folding intermediates is to compare the refolding kinetics measured at different protein concentrations (Silow & Oliveberg, 1997). If there are none, the normalized kinetic trace should be independent of protein concentration. Figure 7 shows such a comparison at protein concentrations from 2.5 to 80 μM; refolding to N is measured after a pH jump from 4.2 to 6.0. The relative amplitude of the slowest kinetic phase increases strongly with concentration (Figure 7(a)). Native protein is formed both in this phase (phase 3) and in the preceding phase: see the discussion of the interrupted refolding experiment in Figure 5. Because the reaction product is the same in phases 2 and 3, but the two reactions have different rates, the starting materials for the two reactions should be different.

As the relative amplitude of phase 3 rises with protein concentration, that of phase 2 falls (Figure 7(b)). Presumably a dimeric or oligomeric intermediate species is responsible. The association constant for its formation may be written:

$$K_n = [\mathrm{I}_n]/[\mathrm{I}]^n \qquad (1)$$

where n is the number of monomeric units in I_n. If I_n forms N in phase 3 while monomeric I forms N in phase 2, and if the reaction between I and I_n reaches equilibrium as N is being formed, then K_n may be written:

$$K_n = (A_3/\Delta F_3)/(A_2/\Delta F_2)^n \qquad (2a)$$

or

$$\log A_3 - n \log A_2 = \text{constant} \qquad (2b)$$

where ΔF_3, ΔF_2 represent the differences between the specific fluorescence coefficients of reactant and product in phases 3 and 2, and A_3, A_2 are the amplitudes of the phases. The plot of log A_3 *versus* log A_2 for the data shown in Figure 7(b) gives a straight line (Figure 7(c)) with slope 1.85, indicating that a dimer ⇌ monomer reaction is rate-limiting for the slow formation of N in phase 3. The

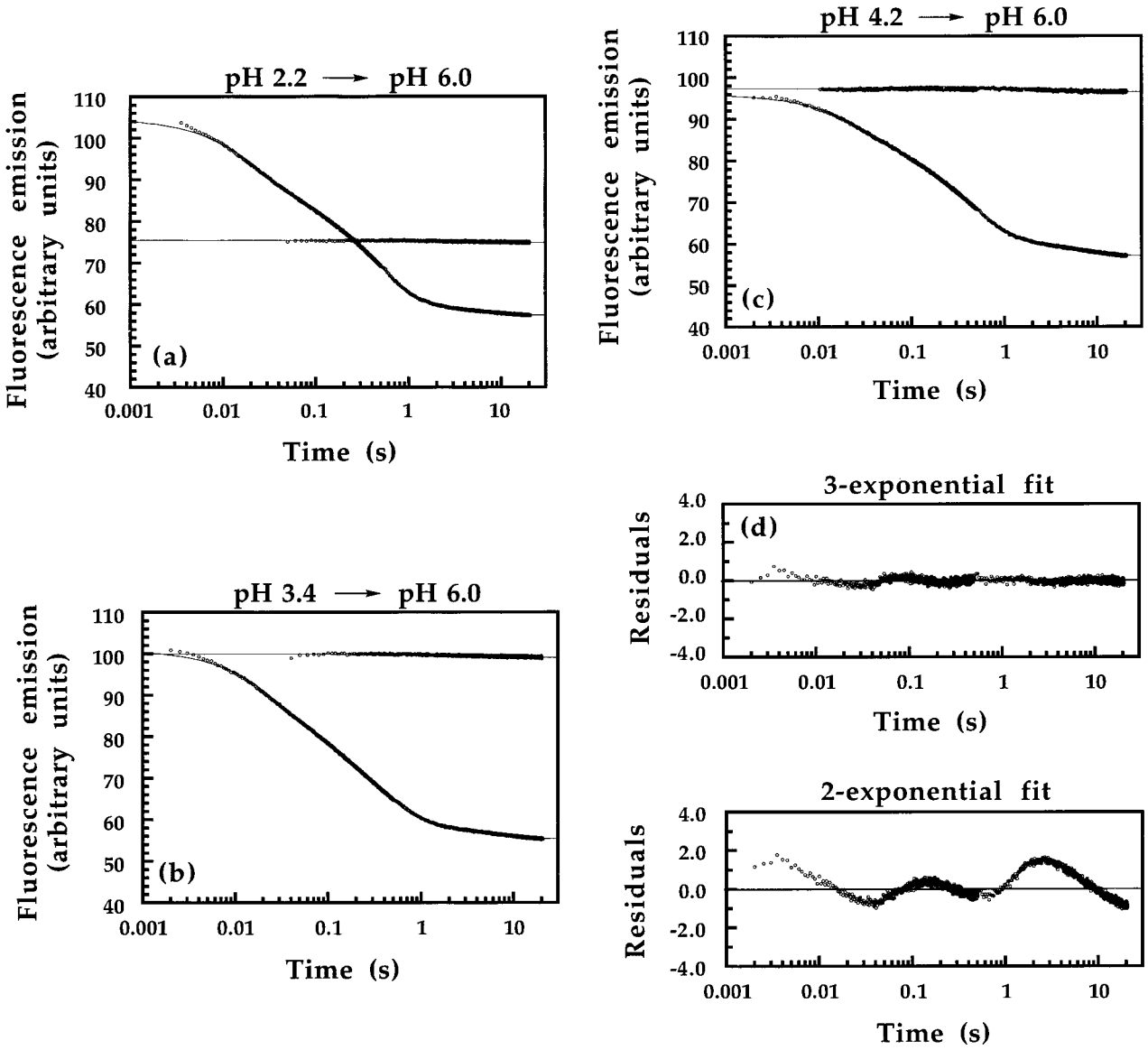

Figure 4. Refolding kinetics at pH 6.0 starting from either the acid-unfolded form (U, pH 2.2) or the intermediate form (I) at pH 3.4 and 4.2. Conditions are the same in all experiments: 2 mM Na citrate/citric acid, 30 mM NaCl, 4.5°C, final protein concentration 10 μM. For each initial pH, two kinetic traces are shown: the almost level one is obtained after a 1:1 dilution of the protein solution at constant initial pH; a small variation of fluorescence with time is observed which is caused by dissociation of dimer. Each line is fitted to a single exponential equation. The second curve shows the refolding kinetics after a pH jump to a fixed final pH 6.0. The curve is fitted to a three-exponential equation (parameters are reported in Table 1). All fluorescence values are normalized so that the initial fluorescence at pH 3.4 (after stopped-flow dilution) is 100. (a) Initial pH 2.2, initial fluorescence at pH 2.2 is 75. (b) Initial pH 3.4, initial fluorescence at pH 3.4 is 100. (c) Initial pH 4.2, initial fluorescence at pH 4.2 is 96. (d) The residuals for fitting either to a three-exponential or a two-exponential equation of the refolding kinetics from pH 4.2 (Figure 4(c)).

experiment below shows directly that a dimeric or oligomeric intermediate dissociates to form native monomer in the slow phase of refolding.

A direct test of whether the rate-limiting step in the formation of I_n is the formation of a dimer is to measure its rate of formation when the concentration of I is varied. The rate of forming I_n should vary as the square of [I] if the rate-limiting step is dimer formation. This test is shown in Figure 8. I is produced from N by an unfolding pH jump from 6.0 to 4.2 and the concentration of I_n (the species formed in phase 3, Figure 6(a)) is measured as a function of time. The experiment is made at protein concentrations of 20, 40, 60 and 80 μM and the results are simulated (see Figure 8) by assuming that I_n is formed in a second-order reaction. The satisfactory fit confirms that I_n is a dimer in the rate determining step of the oligomerization reaction and so I_n may be written I_2. This kinetic experiment only reports on the rate-limiting step of the reaction and does not detect whether dimer can form higher order oligomers. Equilibrium sedimentation experiments will be

Table 1. Analysis of the refolding kinetics of N at pH 6.0 starting from three different initial pH conditions

	pH 2.2	pH 3.4	pH 4.2
A_1	17.3 (36%)	17.1 (38%)	11.5 (29%)
$1/\tau_1$	39.1 s^{-1}	33.6 s^{-1}	30.8 s^{-1}
A_2	26.8 (57%)	24.2 (53%)	23.0 (59%)
$1/\tau_2$	2.2 s^{-1}	3.0 s^{-1}	2.4 s^{-1}
A_3	3.2 (7%)	4.3 (9%)	4.5 (12%)
$1/\tau_3$	0.19 s^{-1}	0.21 s^{-1}	0.17 s^{-1}
F_{burst}	+29.8	+5.4	−3.9
F_{init}	104.7	102.0	96.1
F_{final}	57.4	56.4	57.1

The starting material is either U at pH 2.2 or a mixture of Ia and Ib, in varying proportion, at pH 3.4 and pH 4.2. The refolding kinetics are fitted to a three-exponential equation (see Methods). The kinetic traces are shown in Figure 4. The fluorescence values are normalized such that the initial fluorescence at pH 3.4 is 100. The values in parentheses are the relative amplitudes. F_{init} represents the fluorescence value when all kinetic phases are extrapolated to time zero; it is obtained by adding the tjree amplitudes to the final fluorescence value. The burst phase change (F_{burst}) is the difference between the initial fluorescence measured in the dilution experiments and the initial kinetic fluorescence (F_{init}).

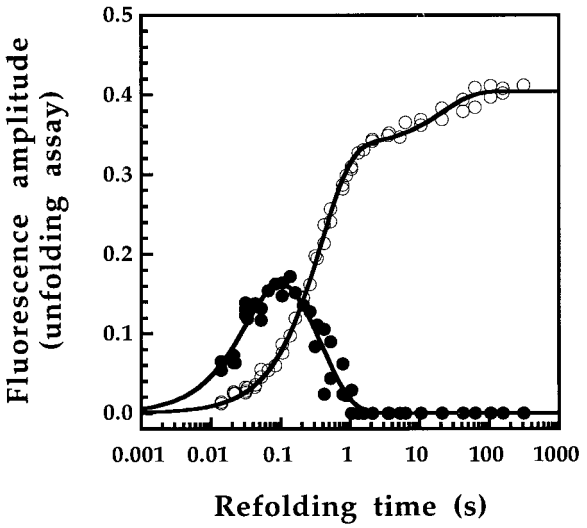

Figure 5. Interrupted refolding experiment (I, pH 4.2 → N, pH 6.0). Refolding is monitored by an assay for the different species present based on their unfolding rates. Unfolding occurs after a pH jump from 6.0 to 4.2. Unfolding occurs in two kinetic phases: (●) fast phase (230 s^{-1}), (○) slow phase (30 s^{-1}); the amplitudes of these two phases are shown *versus* the time when refolding is interrupted. The fast unfolding phase results from the disappearance of a folding intermediate (Ib); the variation of its amplitude with the refolding time shows the transient accumulation of Ib during refolding under native conditions. The slow unfolding phase results from the unfolding of the native protein (N); its amplitude *versus* refolding time represents the kinetics of formation of N during refolding, which occurs in two phases, a major phase followed by a minor phase. Conditions: 2 mM Na citrate/citric acid, 30 mM NaCl, 4.5°C, final protein concentration in refolding 10 μM. The same three kinetic phases are seen as in the direct refolding experiment in Figure 4(c). The filled circles show the accumulation of Ib (phase 1) followed by its disappearance (phase 2). The open circles show the formation of N, either from monomeric Ib (phase 2) or from dimeric I_2 (phase 3). The amplitude and apparent rate constants of the three phases are: (1) formation and disappearance of Ib, $A_1 = 0.234$, $1/\tau_1 = 25.4\ s^{-1}$; $A_2 = -0.234$, $1/\tau_2 = 2.6\ s^{-1}$; (2) formation of N, $A_2 = 0.335$, $1/\tau_2 = 2.6\ s^{-1}$; $A_3 = 0.069$, $1/\tau_3 = 0.05\ s^{-1}$.

necessary to determine if oligomeric species other than dimer are also present at equilibrium. The non-zero intercepts for the amount of I_2 at zero time of unfolding N (Figure 8) may reflect the formation of I_2 during the refolding assay. Because the amount of I_n present at any time is measured as the amount of native monomer formed in the slow phase of a refolding assay, this experiment also confirms that the reaction in phase 3 of refolding is dissociation of oligomer followed by formation of N.

The results in Figure 8 indicate that some I_n should be present at equilibrium at pH 4.2 and the relative amount of I_n should increase with protein concentration. This prediction was tested by stopped-flow experiments in which the protein is diluted while the solvent conditions are held fixed (pH 4.2, standard conditions). A kinetically measurable decrease in fluorescence occurs in the seconds time range ($1/\tau = 0.35\ s^{-1}$) after a fivefold dilution, starting from protein concentrations of 5 to 160 μM (data not shown). The relative amplitude of the reaction increases nonlinearly with protein concentration, as expected. In the time range of the Ia ⇌ Ib reaction (5 to 10 ms), there is no measurable kinetic change in fluorescence caused by dilution, which indicates that Ib is not a dimer of Ia, or *vice versa*.

The transition curve for urea unfolding (pH 4.2, standard conditions) monitored by CD is independent of protein concentration from 10.7 to 53.5 μM (data not shown). This result indicates either that the dimer and monomer are equally stable to urea unfolding or else that the dimer dissociates at low urea concentrations and the transition curve gives the unfolding curve of the monomer. The mean residue ellipticity at 0 M urea does not change measurably between 10 and 80 μM, indicating that dimer and monomer have similar values.

Discussion

Folding pathway of apomyoglobin

We show here that there are two forms of the pH 4 intermediate in equilibrium with each other, separated by a kinetic barrier in the 1 to 10 ms time range, and that both Ia and Ib are formed during the kinetics of folding and unfolding of native apoMb. The basic questions concerning the

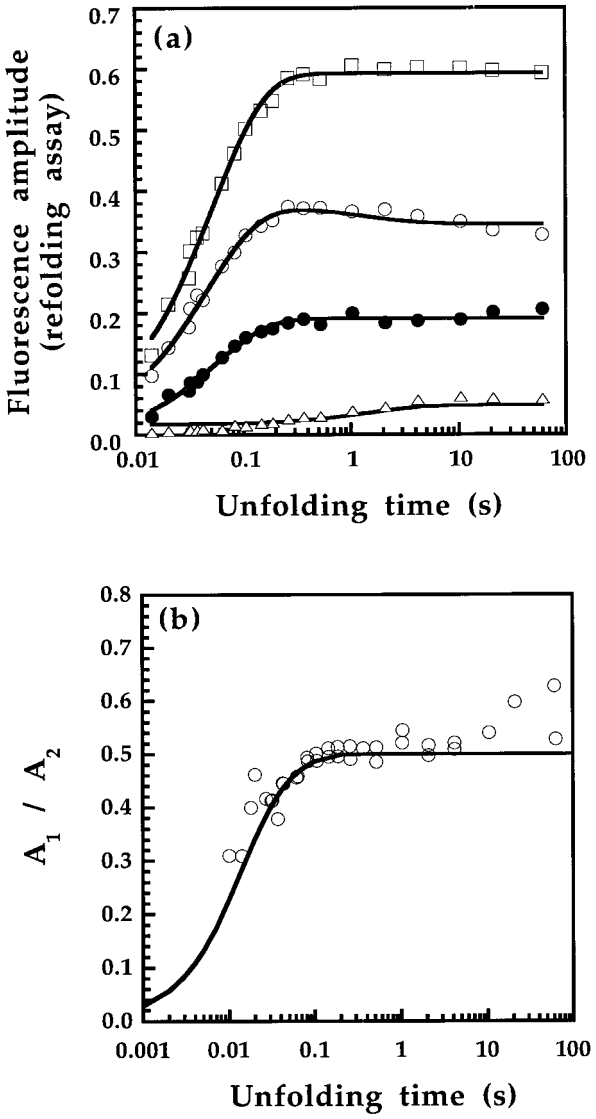

Figure 6. Interrupted unfolding experiment (pH 6.0 → 4.2). Unfolding is monitored by a refolding kinetics assay following a pH jump from 4.2 to 6.0. Refolding occurs in three kinetic phases (see Figures 4(c) and 5). Conditions: 2 mM Na citrate/citric acid, 30 mM NaCl, 4.5°C, final protein concentration in unfolding 10 μM. (a) Amplitudes of the three phases measured from refolding kinetics *versus* time of unfolding. The variation of the amplitudes with unfolding time represents the kinetics of formation of the different species present at equilibrium at pH 4.2: (●) Ib → Ia, (A_1); (○) N → Ib, (A_2); (△) 2I → I_2, (A_3); (□) total amplitude. The lines are drawn as described in Methods. (b) Ratio of the amplitudes A_1/A_2 from Figure 6(a) plotted *versus* time of unfolding. The line is drawn using equation (8) and the following parameters: $\alpha = 0.75$, $k_{12} = 20\ s^{-1}$, $k_{21} = 1\ s^{-1}$, $k_{23} = 80\ s^{-1}$, $k_{32} = 120\ s^{-1}$, $1/\tau_1 = 20.6\ s^{-1}$, $1/\tau_2 = 200.5\ s^{-1}$.

roles of Ia and Ib in the folding pathway of N are: (1) are they on-pathway or off-pathway intermediates, and (2) if they are on-pathway, do they lie on the same pathway or on parallel pathways? In the following discussion, we consider three models, which illustrate contrasting answers to these two questions.

Linear $\quad U \rightleftharpoons Ia \rightleftharpoons Ib \rightleftharpoons N \quad$ (I)

Parallel $\quad$ Ia; U ⇌ Ia, Ia ⇌ N, U ⇌ Ib, Ib ⇌ N $\quad$ (II)

Dead end $\quad$ Ia; U ⇌ Ia, U ⇌ N, U ⇌ Ib $\quad$ (III)

Certain basic properties of the apoMb system should be kept in mind when considering these models. (1) The intermediates Ia, Ib are formed when N unfolds at pH 4.2 (Figure 6) as well as when U refolds to N at pH 6.0 (Figures 4 and 5). (2) The U ⇌ Ia reaction occurs with measurable kinetics during urea-induced unfolding at pH 4.2 (Jamin & Baldwin, 1996; see Figure 3(b)) and follows the two-state model to a first approximation. The Ia ⇌ Ib reaction is at least an order of magnitude slower that the U ⇌ Ia reaction in conditions where both reaction occur. (3) The fractional amount of U is very small in the conditions where the [U]/[Ia] ratio has been measured (pH 4.2). The U ⇌ Ia transition is measured above 1 M urea, where Ib is not populated (Figure 3(b)) and an equilibrium constant for the U ⇌ Ia reaction close to 100 is calculated from the extrapolation to 0 M urea of the folding and unfolding rate constants measured in the transition region. The additional presence of Ib at 0 M urea means that the [U]/([Ia] + [Ib]) ratio is smaller than the [U]/[Ia] ratio and that there is less than 1% U present in these conditions. The equilibrium constant for the Ia ⇌ Ib reaction has not yet been measured. (4) The intermediates Ia, Ib are rapidly interconvertible ($\tau = 6$ ms) at pH 4.2, where they are practically the only species present.

Rapid interconversion between Ia and Ib at pH 4.2 means that models II and III are unlikely because the amounts of U and N are very small at equilibrium at pH 4.2 and, in these models, Ia is transformed into Ib only by first unfolding to U (proposed in both models) or by first refolding to N (in model II). Model III has been tested directly, using an unfolding rate assay that is specific for N, by assaying for "fast track" U → N folding that bypasses Ib and Ia. No fast-track folding has been found, either after the pH jump pH 4.2 → 6.0 (Figure 5) or after the pH jump pH 2.2 → 6.0 (data not shown), for which the starting material is U. Model II has been tested directly by asking if the [Ia]/[Ib] ratio is independent of the time of unfold-

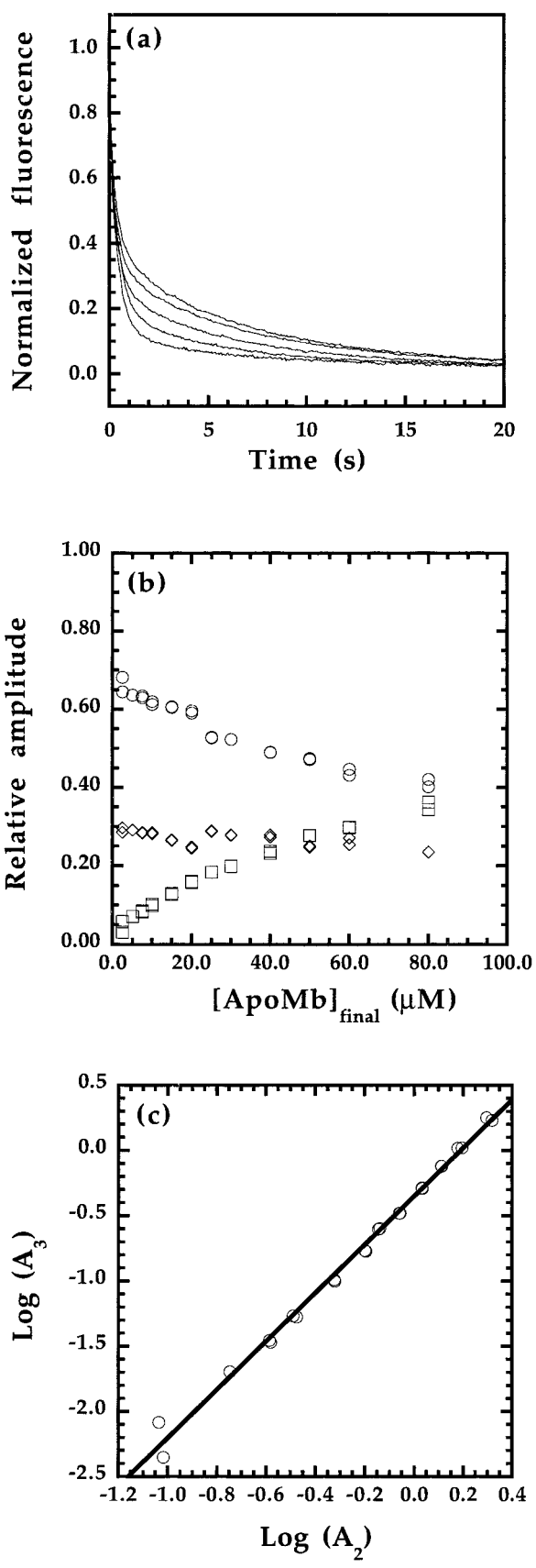

Figure 7. Dependence on protein concentration of refolding kinetics at pH 6.0. Refolding is initiated by a pH jump from 4.2 to 6.0 at final protein concentrations varying from 2.5 to 80 μM. (a) Increase in the relative amplitude of phase 3 with [apoMb]: 10 μM (lowest curve), 20, 40, 60 and 80 μM (upper curve). (b) Relative amplitudes of the three phases *versus* [apoMb]: (◇) phase 1 (fast), (○) phase 2 (middle), (□) phase 3 (slow). (c) Logarithmic plot of the amplitude of the slow phase *versus* the amplitude of the middle phase. The line has a slope of 1.85 and an intercept of −0.35.

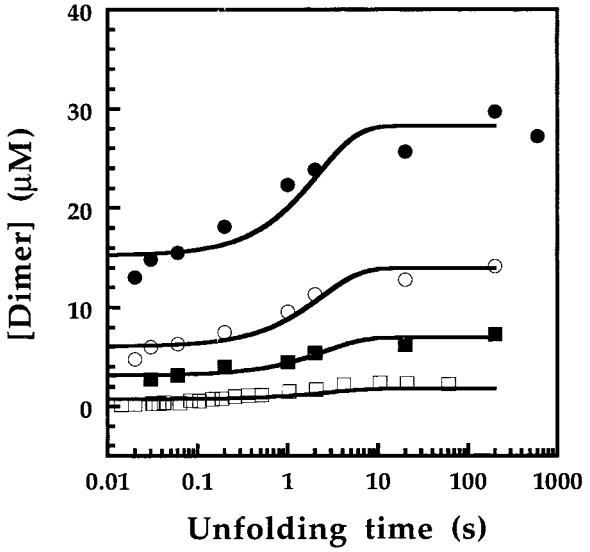

Figure 8. Kinetics of dimer formation at pH 4.2. Monomeric intermediate is formed at pH 4.2 by unfolding N after a pH jump from 6.0 to 4.2; the intermediate then forms dimers. The amount of dimer formed is plotted *versus* unfolding time for four different final protein concentrations: 80 μM (●), 60 μM (○), 40 μM (■), and 20 μM (□). See Methods for the procedures used to analyze and simulate the results. The amount of dimer is measured as the amount of native monomer formed in phase 3 of the refolding assay. The parameters used in the simulation are as follows: $k_{on} = 930\ M^{-1}\ s^{-1}$ and $k_{off} = 0.32\ s^{-1}$.

ing, as expected for a parallel pathway model. Figure 6 shows that the ratio [Ia]/[Ib] increases with unfolding time at very early times (Figure 6(b)) and then reaches a constant value, when Ia and Ib equilibrate following the slower N → Ib reaction. Thus, the experiment in Figure 6 contradicts model II. On the other hand, all our results are consistent with the linear model I.

The possibility that stable folding intermediates represent dead-end intermediates has been given prominence by simulations of folding (see reviews by Eaton *et al*., 1996; Dill & Chan, 1997). The argument is that, if multiple pathways are available for folding, then folding will go by the fastest route and it should bypass stable intermediates, which represent kinetic traps. Experiments have demonstrated competing pathways for folding and unfolding but these experiments show typically only two pathways, not multiple pathways (dihydrofolate reductase, Jennings *et al*. 1993; barstar,

Shastry & Udgaonkar, 1995; Zaidi *et al*., 1997; hen lysozyme, Kiefhaber, 1995; Wildegger & Kiefhaber, 1997; Matagne *et al*., 1997). These results provide strong support for the view that alternative pathways are available for folding and unfolding. NMR-hydrogen exchange experiments on hen lysozyme had suggested earlier that multiple pathways are available (Radford *et al*., 1992). Nevertheless, the results for some carefully studied systems support the view, as we find here for apomyoglobin, that folding intermediates are frequently productive and consequently that the term ''folding pathway'' is meaningful: see the barnase results of Oliveberg & Fersht (1996), the ribonuclease H results of Chamberlain & Marqusee (1997), the barstar results of Zaidi *et al*. (1997), the hen lysozyme results of Wildegger & Kiefhaber (1997), and the interleukin-1β results of Heidary *et al*. (1997). Resolution of this controversy probably awaits the time when folding simulations explore the properties of these stable folding intermediates. From our results, we expect that appropriate folding simulations will then show stable, on-pathway folding intermediates.

The dimeric intermediate reported here is evidently not an obligatory intermediate on the folding pathway of apoMb because at low protein concentrations (<10 μM) it is formed slowly relative to the monomeric folding pathway, and the rate of forming N on the monomeric pathway is independent of protein concentration. Only a small fraction of N is formed *via* the dimeric intermediate at these low concentrations. At higher concentrations, however, dimer formation becomes a strongly competing pathway.

Dimer formation is undoubtedly strongly dependent on anion conditions and temperature, like the Ia ⇌ Ib reaction. We have studied dimer formation so far only in the conditions shown in Figures 7 and 8.

A dimeric folding intermediate of horse heart apoMb was discovered by Eliezer *et al*. (1993), in a small angle X-ray scattering study of the refolding kinetics of native apoMb. They used a protein concentration of about 500 μM, as compared to the standard concentration of 10 μM used here for kinetic studies, and the amount of dimer formed increases rapidly with increasing protein concentration. This result is consistent with our finding that a dimer ⇌ monomer reaction is rate-limiting for the dissociation of oligomeric intermediate species under native conditions.

Relation to other studies of apomyoglobin folding intermediates

Apomyoglobin presents an unusual example of a protein showing both a low-salt folding intermediate at pH 4, which becomes unfolded on lowering the pH to 2 (Griko *et al*., 1988), and a high-salt folding intermediate at pH 2, whose existence requires either mM concentrations of a stabilizing anion, such as sulfate, perchlorate or trichloroacetate, or a high concentration (0.5 M) of a weakly stabilizing anion, such as chloride (Goto *et al*., 1990). The relation between the pH 4 low-salt and pH 2 high-salt intermediates remains to be worked out.

The discovery that the pH 4 folding intermediate itself exists in two different forms (Ia, Ib), separated by a kinetic barrier, suggests that a similar kinetic barrier might exist between Ia and the pH 2 high-salt intermediate. The observation that Ia is converted to a structurally different form in which the B helix becomes stable in the presence of 20 mM TCA, pH 4.2 (Loh *et al*., 1995), shows that important structural differences can exist between the different folding intermediates. Unfortunately, because of the strong fluorescence photobleaching of I in the presence of TCA, we have not been able to test whether Ib is the same as the TCA form of Loh *et al*. (1995). Studies of the pH 2 high-salt intermediate have typically been made with horse heart rather than sperm whale apomyoglobin (see Nishii *et al*., 1995), but the properties of the two proteins appear to be similar. A study by Rischel *et al*. (1996) of fluorescence energy transfer between a modified Tyr residue in the H helix and the two Trp residues of the A helix confirms the presence of a close contact between the A and H helices in the pH 4 intermediate of horse heart apoMb.

The Ia ⇌ Ib reaction, which occurs with a relaxation time of 30 ms at pH 6.0 (Table 1), was not observed in the study by Jennings & Wright (1993), who used CD (222 nm) and NMR-hydrogen exchange as probes of folding. It is possible that both probes are insensitive to the Ia ⇌ Ib reaction, which is monitored here by tryptophan fluorescence. Alternatively, since both the signal-to-noise ratio and the mixing dead time are more favorable for resolving the U ⇌ Ia and Ia ⇌ Ib reactions by fluorescence than by CD, and because hydrogen exchange is not monitored continuously during folding, their experiments may not have resolved the U ⇌ Ia reaction from the Ia ⇌ Ib reaction. We repeated the refolding experiment of Jennings & Wright (1993), which has 0.8 M urea in 10 mM sodium acetate, pH 6.0, in the final folding conditions, and found that the Ia → Ib reaction is still observed and is somewhat slower ($1/\tau_1 = 20\ s^{-1}$, $1/\tau_2 = 2.5\ s^{-1}$, $1/\tau_3 = 0.25\ s^{-1}$). The Ia ⇌ Ib reaction observed during the formation of N at pH 6.0 has a significantly smaller amplitude in 10 mM acetate (as in the study of Jennings & Wright, 1993) than in the citrate results shown here (see Figure 4; our acetate data are not shown).

Jennings & Wright observed by hydrogen exchange a later folding intermediate (time range one second) which we do not observe here by tryptophan fluorescence. The entire B helix becomes stable in this intermediate, whereas a few protons near the N terminus of the B helix are stable in their initial intermediate, which is formed within the stopped-flow mixing time.

Whether the pH 4 intermediate has a well-defined structure that is formed cooperatively, or

whether its structure changes continuously as the conditions are varied, is the subject of conflicting reports: see Griko & Privalov (1994), Gast *et al.* (1994), Nishii *et al.* (1995), Kay & Baldwin (1996), Jamin & Baldwin (1996), Luo *et al.* (1997). The result reported here, namely that there are two forms of the pH 4 intermediate separated by a kinetic barrier, adds a new twist to the problem. The changes in spectral properties with temperature (Gast *et al.*, 1994; Nishii *et al.*, 1995) may be the result of interconversion between two well-defined forms rather than a continuous variation in structure with temperature.

Time-resolved infrared spectroscopy and laser-induced temperature jumps were used by Gilmanshin *et al.* (1997) to monitor fast changes in the helical backbone as native apoMb (horse heart) unfolds. They found evidence for helix folding/unfolding in two widely separated time ranges (10^{-7} seconds and 10^{-3} seconds), which they attributed to solvent exposed and native helices, respectively. More recently, they made thermal unfolding experiments with the I form (pH 3.15, 0.15 M NaCl), which shows a similar series of amide I bands corresponding to native helices, solvent exposed helices and random coil (R. Gilmanshin *et al.*, unpublished results). Only the solvent exposed helices undergo unfolding in the temperature range they study. A fast fluorescence change (10^{-5} seconds) following a laser-induced temperature jump was observed by Ballew *et al.* (1996) at low temperatures, where native apoMb undergoes cold denaturation. Thus, fluorescence is able to monitor very early events in folding/unfolding of native apoMb. These reactions are too fast to be detected by us.

Materials and Methods

Unless otherwise stated, experiments were performed at 4.5°C in 2 mM Na citrate/citric acid and 30 mM NaCl. pH values below 3.5 were obtained by adding small amounts of hydrochloric acid.

Protein expression and purification

A synthetic gene for sperm whale myoglobin carried by the pMbT7 plasmid (Loh *et al.*, 1995) was expressed in *Escherichia coli* strain BL21. The protein was purified as described (Loh *et al.*, 1995). The heme was removed by acid-acetone precipitation (Fanelli *et al.*, 1958). The protein was >95% pure and homogenous as judged by PAGE in the presence of SDS. No irreversible aggregate was detected upon elution of the protein on a superose 12 column (Pharmacia, Uppsala, Sweden) under standard conditions (100 mM Na phosphate, pH 6.0) (Uversky, 1993). A strong nonspecific ionic interaction with the gel precludes the use of the low ionic strength buffer used here for the kinetic studies: in 2 mM Na citrate, 30 mM NaCl, pH 6.0, apomyoglobin elutes after the salts. Spinning the solution at 2000 *g* did not affect either the global fluorescence change or the relative amplitudes of the three observed refolding phases.

Protein concentrations were determined by absorbance in 6.0 M GdmCl (20 mM Na phosphate, pH 6.5) as described (Edelhoch, 1967), using $\varepsilon_{280\ nm} = 15{,}200\ M^{-1}\ cm^{-1}$ and $\varepsilon_{288\ nm} = 10{,}800\ M^{-1}\ cm^{-1}$.

Fluorescence and CD equilibrium measurements

Protein concentrations in equilibrium studies were between 1 and 10 μM. Fluorescence data were collected on a SLM AB2 spectrofluorimeter using a 0.4 cm × 1.0 cm cuvette thermostatted at 4.5°C. The excitation was set at 288 nm with a bandpass of 1 nm. The emission was recorded between 320 and 380 nm with a bandpass of 8 nm. The integral of the spectrum between 320 and 380 nm was normalized to a value of 100 at pH 3.4.

Circular dichroism was measured at 222 nm on an AVIV 62A DS spectropolarimeter using a 1 mm cuvette and a thermostatted sample holder set at 4.5°C.

Fluorescence stopped-flow studies

Data were collected on a sequential mixing stopped-flow instrument (model DX.17 MV) from Applied Photophysics. The excitation was set at 288 nm (slit width 1 mm) and an optical cut-off filter (50% transmittance at 320 nm) was used to detect the emission. The pathlength of the cell was 1 cm. For each experiment, between four and 12 kinetic traces were averaged. Fluorescence values were normalized such that the initial fluorescence of the protein at pH 3.4 is 100. The stopped-flow mixing time was determined by measuring a standard second-order reaction (Tonomura *et al.*, 1978) at varying initial concentrations. The kinetic traces obtained in the stopped-flow experiments were fitted to single, double or triple exponentials with the software provided by Applied Photophysics.

Fluorescence is proportional to concentration over only a limited range of absorbance (Lakowicz, 1986). In experiments where protein concentration was varied, kinetic amplitudes were corrected for the inner filter effect by using the equation:

$$F = F_{obs} \exp(2.303 \varepsilon_{288\ nm} CL/2) \quad (3)$$

where F_{obs} is the measured fluorescence, $\varepsilon_{288\ nm}$ is the extinction coefficient for the protein at 288 nm (10800 M^{-1} cm^{-1} in 2 mM Na citrate, 30 mM NaCl, pH 4.2), C is the protein concentration and L is the pathlength of the cell.

Unimolecular reactions

The kinetic analysis of reversible unimolecular reactions has been described in detail by Szabo (1969). Analytical solutions can be found for the sequential three-state mechanism :

$$A \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} B \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} C \quad (4)$$

This mechanism gives two macroscopic rate constants $1/\tau_1$ and $1/\tau_2$, which depend only on the microscopic rate constants:

$$1/\tau_{1,2} = 0.5(b \pm \sqrt{b^2 - 4c}) \quad (5)$$

where $b = k_{12} + k_{21} + k_{23} + k_{32}$ and $c = k_{12}(k_{23} + k_{32}) + k_{21}k_{32}$.

Equations describing the variation of concentrations of the three different species with time ($[A]_t$, $[B]_t$, $[C]_t$) can also be found, which depend on both the microscopic rate constants and the initial concentrations. For example, when the reaction is started from 100% *A*, the following equations were derived with the help of Mathematica (Wolfram Research, Inc., Champaign, IL, USA):

$$[A]_t = [A]_0\left(\frac{k_{21}k_{32}}{(1/\tau_1)(1/\tau_2)} - \frac{k_{12}(k_{23}+k_{32}-1/\tau_1)}{1/\tau_1(1/\tau_1-1/\tau_2)}\exp(-t/\tau_1) + \frac{k_{12}(k_{23}+k_{32}-1/\tau_2)}{1/\tau_2(1/\tau_1-1/\tau_2)}\exp(-t/\tau_2)\right) \quad (6a)$$

$$[B]_t = [A]_0\left(\frac{k_{12}k_{32}}{(1/\tau_1)(1/\tau_2)} + \frac{k_{12}(k_{32}-1/\tau_1)}{1/\tau_1(1/\tau_1-1/\tau_2)}\exp(-t/\tau_1) - \frac{k_{12}(k_{32}-1/\tau_2)}{1/\tau_2(1/\tau_1-1/\tau_2)}\exp(-t/\tau_2)\right) \quad (6b)$$

$$[C]_t = [A]_0\left(\frac{k_{12}k_{23}}{(1/\tau_1)(1/\tau_2)} + \frac{k_{12}k_{23}}{1/\tau_1(1/\tau_1-1/\tau_2)}\exp(-t/\tau_1) - \frac{k_{12}k_{23}}{1/\tau_2(1/\tau_1-1/\tau_2)}\exp(-t/\tau_2)\right) \quad (6c)$$

Single jump experiments

The reactions were initiated by rapidly diluting a protein solution at a given pH with an equal volume of buffer solution, chosen to give the desired pH after mixing. In each experiment, the final pH was measured independently by mixing equal volumes of the different solutions. In the protein dilution experiments in constant conditions, the pH of the buffer solution was equal to the pH of the protein solution.

Interrupted refolding

To determine the rate of formation of the native protein assayed by its unfolding rate (Schmid, 1983), we used a double-jump stopped-flow experiment. In the first step, the reaction was initiated by 1:1 dilution of the pH 4.2 intermediate into refolding buffer to bring the pH to 6.0. After a variable time of refolding, the pH was transferred back to 4.2 and the unfolding reaction was monitored by Trp fluorescence. After a long time (two seconds) of refolding, the time course of unfolding (fluorescence increase) was monoexponential with a rate constant of 30 s^{-1}. After shorter times of refolding, a second faster phase was observed with a rate constant of 230 s^{-1}. In a control experiment, the native protein at equilibrium at pH 6.0 was jumped to pH 4.2 and monoexponential unfolding was observed with a rate constant of 25 s^{-1}. The unfolding amplitudes, measured at various refolding times, yield the kinetics for transient accumulation of the intermediate Ib and for the formation of N. The accumulation of Ib and the formation of N can each be fitted to a two-exponential equation. The rate constants obtained from the analysis of these interrupted refolding kinetics match with the rate constants of the three refolding phases observed directly after a single jump from pH 4.2 to pH 6.0. It is important to note that the rate of disappearance of Ib is the same as the fast rate of formation of N.

Interrupted unfolding

To determine the kinetics of formation of the different species after a pH jump from 6.0 to 4.2, we used a refolding assay. First, the protein at pH 6.0 was diluted 1:1 to pH 4.2 and, after a variable time of unfolding, a second 1:1 dilution was performed to bring the pH back to 6.0. The refolding kinetics were monitored by Trp fluorescence. Three phases with similar rate constants were observed after the earliest times of unfolding. Figure 6(a) represents the amplitudes of the three phases as a function of unfolding time.

The unfolding reaction can be represented by the following model:

$$\mathrm{N} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \mathrm{Ib} \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} \mathrm{Ia} \quad (7)$$

If the reaction is started from 100% N, the time courses for Ia and Ib provide a value for the ratio $([\mathrm{Ia}]/[\mathrm{Ib}])_t$. To convert the ratio of fluorescence amplitudes into a ratio of concentrations: (1) the two steps N → Ib and Ib → Ia must occur on different time scales such that their rates are not coupled; and (2) the change in the specific fluorescence coefficients associated with each reaction must be known. An equation which gives A_1/A_2 *versus* unfolding time was obtained from equations (6b) and (6c):

$$\frac{A_1}{A_2} = \alpha\left\{\frac{k_{23}(1/\tau_1-1/\tau_2)+k_{23}(1/\tau_2)\exp(-t/\tau_1)-k_{23}(1/\tau_1)\exp(-t/\tau_2)}{k_{32}(1/\tau_1-1/\tau_2)+(k_{32}-1/\tau_1)(1/\tau_2)\exp[(-t/\tau_1)-(k_{32}-1/\tau_2)(1/\tau_1)\exp(-t/\tau_2)}\right\} \quad (8)$$

where $1/\tau_1$ and $1/\tau_2$ are given by equation (5), and $\alpha = \Delta F_1/\Delta F_2$. Thus, the value of α depends on the specific fluorescence coefficients for Ia, Ib and N. The fluorescence coefficient for N is known from the fluorescence at the end of the refolding kinetics at pH 6.0 and a good estimate of the coefficient for Ia is obtained from the burst phase fluorescence in the pH 2.2 → pH 6.0 kinetics (Figure 4(a)). The value for Ia is determined by extrapolating the measured kinetic phases to time zero. The specific fluorescence of Ib is not known because conditions where Ib is 100% populated cannot be found. From the available information, we estimated a value of $\alpha = 0.75$ and used it to fit the data. The line decribing A_1/A_2 in Figure 6(b) was checked by numerical simulation using SIMFIT (Holzhütter & Colosimo, 1990).

In Figure 6(a), the variation of A_1 and of the total amplitude ($A_1 + A_2 + A_3$) with unfolding time are fitted to a single exponential equation:

$$F = A(1-\exp(-t/\tau)) \quad (9)$$

with the following parameters: fast phase amplitude: $A_1 = 0.19$, $1/\tau_1 = 17.0$ s^{-1}; total amplitude: $A = 0.55$, $1/\tau = 17.8$ s^{-1}. The variation of A_3 with unfolding time, which represents the kinetics of dimer formation at pH 4.2, is fitted to (equation (13) as described in the section below with $k_{on} = 1620$ M^{-1} s^{-1}, $k_{off} = 0.55$ s^{-1}, intercept = 0.0168. The variation of A_2 with unfolding time is

fitted to:

$$F = F_{\text{final}} - A_2 \exp(-t/\tau_2) - [\text{I}_2]_t \quad (10)$$

where $[\text{I}_2]_t$ is given by equation (13) using the same parameters as in the analysis of A_3 *versus* unfolding time. The following parameters are found: $A_2 = 0.35$, $1/\tau_2 = 20.1\ \text{s}^{-1}$, $F_{\text{final}} = 0.39$.

Interrupted unfolding: dimer formation

The interrupted unfolding experiment was performed at four different protein concentrations. The amplitudes were corrected for the inner filter effect using equation (3), with $L = 1$ cm and $\varepsilon_{288\ \text{nm}} = 10{,}800\ \text{M}^{-1}\ \text{cm}^{-1}$. Then the fluorescence amplitudes were converted into dimer concentration using the following information: (1) from the interrupted refolding experiment of Figure 5, we know that at pH 4.2, in a 40 μM protein solution, 17% of the protein molecules form dimers, (2) from the refolding kinetics (pH 4.2 → pH 6.0), we know the fluorescence change that occurs when dimers form monomeric N. Thus, we can compute the change in specific fluorescence coefficients when N unfolds to intermediate which then forms dimers.

The dimerization reaction can be represented by:

$$2\text{I} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \text{I}_2 \quad (11)$$

This model gives

$$\frac{\text{d}[\text{I}]}{\text{d}t} = -k_{\text{on}}[\text{I}]^2 + k_{\text{off}}[\text{I}_2] \quad (12)$$

Integration of equation (12) between time 0 and t, yields an equation for the time course of I_2 (see for example Milla & Sauer, 1994). Only monomer is present at 0 time because unfolding starts from N and N is assumed not to form dimers. In these conditions, when the reaction is initiated from 100% monomer, $[\text{I}]_0$ can be equated to $[\text{I}_\text{T}]$ and the time course of $[\text{I}_2]_t$ is given by the rate constants k_{on} and k_{off} and the global concentration $[\text{I}_\text{T}]$:

$$[\text{I}_2]_t = ([\text{I}_\text{T}] - [\text{I}]_t) + \text{Intercept} \quad (13)$$

where $[\text{I}_2]_t$ is the concentration of dimer at time t (expressed in amount of monomer), $[\text{I}_\text{T}]$ is the global concentration of protein and $[\text{I}]_t$ is the concentration of monomeric species at time t, given by

$$[\text{I}]_t = \frac{(bn + mn - b + n)}{(1 - n)} \cdot \frac{[\text{I}_\text{T}]}{2} \quad (14)$$

where

$$b = \frac{k_{\text{off}}}{k_{\text{on}}[\text{I}_\text{T}]}, \quad m = \sqrt{b^2 - 4b}$$

$$n = \frac{2 + b - m}{2 + b + m} \exp(-mk_{\text{on}}[\text{I}_\text{T}]t)$$

The intercept parameter is introduced in equation (13) to account for the fact that the curves do not extrapolate to 0 at time 0. The meaning of this parameter is discussed in the paper.

The four curves in Fig 8 were fitted jointly to (equation (13) using SigmaPlot (Jandel Software, San Rafael, CA, USA). The lines in Figure 8 are drawn with (equation (13) and the following par ameters: $k_{\text{off}} = 0.32\ \text{s}^{-1}$, $k_{\text{on}} = 930\ \text{M}^{-1}\ \text{s}^{-1}$, intercept (20 μM) = 0.74 μM, intercept (40 μM) = 3.14 μM, intercept (60 μM) = 6.06 μM, intercept (80 μM) = 15.22 μM. Because the monomeric I is formed from N during unfolding, we also simulated the data in a procedure where the amount of $[\text{I}_\text{T}]$ is given by the time course of the N → Ib reaction. No significant difference was found, because formation of I from N is much faster than dimer formation.

Urea-induced unfolding transition

Urea was from US Biochemicals (Ultrapure grade) and the concentration of the stock solution was determined by measuring the refractive index (Pace *et al.*, 1986).

The urea-induced unfolding transitions monitored by CD and fluorescence and also the folding and unfolding kinetics of these transitions are analyzed with the sequential model:

$$\text{U} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \text{Ia} \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} \text{Ib} \quad (15)$$

In our system, we assume that the logarithm of each microscopic rate constant depends linearly on urea concentration (Aune & Tanford, 1969) and can be represented by the following equation:

$$k_{ij} = k_{ij}(\text{H}_2\text{O}) \exp\left(\frac{m_{ij}[\text{urea}]}{RT}\right) \quad (16)$$

where m_{ij} represents the dependence on urea concentration of the rate constant k_{ij} and $k_{ij}(\text{H}_2\text{O})$ is the rate constant in the absence of denaturant. The microscopic rate constants (k_{ij}) in the folding direction decrease with increasing urea concentration, whereas they increase in the unfolding direction.

For the analysis of the equilibrium transition curves, we use the procedure of Santoro & Bolen (1988). Variations of tryptophan and tyrosine fluorescence with urea concentration have been reported (Schmid, 1989) and variations of both fluorescence and CD of the unfolded form (U) with urea concentration are clearly observed in the post-transition region (Figure 3(a)). Because similar changes are also expected to occur with the other forms of the protein (Ia and Ib), we assumed simple linear baselines for each species: for fluorescence, the same slope (b_F) is given to the three baselines; for CD, a different slope is given to the baseline of U (b_U) and to the baselines of Ia and Ib (b_I). The transitions are fitted to the following equations:

$$F = \frac{(F_\text{U} + b_\text{F}[\text{urea}]) + (F_\text{Ia} + b_\text{F}[\text{urea}])K_{12} + (F_\text{Ib} + b_\text{F}[\text{urea}])K_{12}K_{23}}{1 + K_{12} + K_{12}K_{23}} \quad (17)$$

$$CD = \frac{(CD_\text{U} + b_\text{U}[\text{urea}]) + (CD_\text{Ia} + b_\text{I}[\text{urea}])K_{12} + (CD_\text{Ib} + b_\text{I}[\text{urea}])K_{12}K_{23}}{1 + K_{12} + K_{12}K_{23}} \quad (18)$$

where

$$K_{12} = \frac{k_{12}}{k_{21}} \quad \text{and} \quad K_{23} = \frac{k_{23}}{k_{32}}$$

The macroscopic rate constants $1/\tau_1$ and $1/\tau_2$ are obtained by introducing the corresponding equation (16) for each microscopic rate constant into equation (5). In Figure 3, a combined analysis of the kinetic curve ($1/\tau_2$ *versus* urea molarity) and the equilibrium transition curves (fluorescence and CD *versus* urea molarity) were fitted jointly using SigmaPlot. The lines in Figure 3(a) and (b) are drawn with the following parameters: $k_{12}(H_2O) = 2620\ s^{-1}$, $k_{21}(H_2O) = 43\ s^{-1}$, $k_{23}(H_2O) = 120\ s^{-1}$, $k_{32}(H_2O) = 80\ s^{-1}$, $m_{12} = 0.77$ kcal mol^{-1} M^{-1}, $m_{21} = 0.40$ kcal mol^{-1} M^{-1}, $m_{23} = 0$ kcal mol^{-1} M^{-1}, $m_{32} = 1.29$ kcal mol^{-1} M^{-1}, $F_U = 59$, $F_{Ia} = 113$, $F_{Ib} = 88$, $b_F = 2.8\ M^{-1}$, $CD_U = -3440$ deg cm^2 dmol^{-1}, $CD_{Ia} = CD_{Ib} = -15{,}200$ deg cm^2 dmol^{-1}, $b_U = 436$ deg cm^2 dmol^{-1} M^{-1}, $b_I = 1740$ deg cm^2 dmol^{-1} M^{-1}.

Acknowledgments

We thank Bernhard Geierstanger and Jon Goldberg for discussion of these results, and we thank Dr R. Gilmanshin for a preprint of his manuscript (Gilmanshin *et al.*, 1977b). This research was supported by NIH grant GM19988.

References

Aune, K. C. & Tanford, C. (1969). Thermodynamics of the denaturation of lyzozyme by guanidine hydrochloride. II. Dependence on denaturant concentration at 25°C. *Biochemistry,* **11**, 4586–4590.

Ballew, R. M., Sabelko, J. & Gruebele, M. (1996). Direct observation of fast protein folding: the initial collapse of apomyoglobin. *Proc. Natl Acad. Sci. USA,* **93**, 5759–5764.

Chamberlain, A. K. & Marqusee, S. (1997). Touring the landscapes: partially folded proteins examined by hydrogen exchange. *Curr. Biol.* **5**, 859–863.

Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19.

Eaton, W. A., Thompson, P. A., Chan, C-K., Hagen, S. J. & Hofrichter, J. (1996). Fast events in protein folding. *Structure,* **4**, 1133–1139.

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

Eliezer, D., Chiba, K., Tsuruta, H., Doniach, S., Hodgson, K. O. & Kihara, H. (1993)). Evidence of an associative intermediate on the myoglobin refolding pathway. *Biophys. J.* **65**, 912–917.

Fanelli, A. R., Antonini, E. & Caputo, A. (1958). Studies of the structure of hemoglobin. I. Physicochemical properties of human globin. *Biochim. Biophys. Acta,* **30**, 608–615.

Gast, K., Damaschun, H., Misselwitz, R., Müller-Frohne, M., Zirwer, D. & Damaschun, G. (1994). Compactness of protein molten globules: temperature-induced structural changes of the apomyoglobin folding intermediate. *Eur. Biophys. J.* **23**, 297–305.

Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H. & dyers, R. B. (1997). Fast events in protein folding: relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl Acad. Sci. USA,* **94**, 3709–3713.

Goto, Y., Takahashi, N. & Fink, A. L. (1990). Mechanism of acid-induced folding of proteins. *Biochemistry,* **29**, 3480–3488.

Griko, Y. V. & Privalov, P. L. (1994). Thermodynamic puzzle of apomyoglobin unfolding. *J. Mol. Biol.* **235**, 1318–1325.

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

Heidary, D. K., Gross, L. A., Roy, M. & Jennings, P. A. (1997). Evidence for an obligatory intermediate in the folding of interleukin 1-β. *Nature Struct. Biol.* **4**, 725–731.

Holzhütter, H. G. & Colosimo, A. (1990). SIMFIT: a microcomputer software-toolkit for modelistic studies in biochemistry. *CABIOS,* **6**, 23–28.

Hughson, F. M., Wright, P. E. & Baldwin, R. L. (1990). Structural characterization of a partly folded apomyoglobin intermediate. *Science,* **249**, 1544–1548.

Jamin, M. & Baldwin, R. L. (1996). Refolding and unfolding kinetics of the equilibrium folding intermediate of apomyoglobin. *Nature Struct. Biol.* **3**, 613–618.

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.

Jennings, P. A., Finn, B. E., Jones, B. E. & Matthews, C. R. (1993). A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. *Biochemistry,* **32**, 3783–3789.

Kay, M. S. & Baldwin, R. L. (1996). Packing interactions in the apomyglobin folding intermediate. *Nature Struct. Biol.* **3**, 439–445.

Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl Acad. Sci. USA,* **92**, 9029–9033.

Lakowicz, J. R. (1986). *Principles of Fluorescence Spectroscopy,* Plenum Press, NY.

Loh, S. N., Kay, M. S. & Baldwin, R. L. (1995). Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway. *Proc. Natl Acad. Sci. USA,* **92**, 5446–5450.

Luo, Y., Kay, M. S. & Baldwin, R. L. (1997). Cooperativity of folding of the apomyoglobin pH 4 intermediate studied by glycine and proline mutations. *Nature Struct. Biol.* **4**, 925–930.

Matagne, A., Radford, S. E. & Dobson, C. M. (1997). Fast and slow tracks in lysozyme folding: insight into the role of domains in the folding process. *J. Mol. Biol.* **267**, 1068–1074.

Milla, M. E. & Sauer, R. T. (1994). P22 arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry,* **33**, 1125–1133.

Nishii, I., Kataoka, M. & Goto, Y. (1995). Thermodynamic stability of the molten globule states of apomyoglobin. *J. Mol. Biol.* **250**, 223–238.

Oliveberg, M. & Fersht, A. R. (1996). A new approach to the study of transient protein conformations: the formation of a semiburied salt limk in the folding pathway of barnase. *Biochemistry,* **35**, 6795–6805.

Pace, C. N., Shirley, B. A. & Thomson, J. A. (1989). Measuring the conformational stability of a protein. In *Protein Structure: a Practical Approach* (Creighton, T. E., ed.), pp. 311–330, IRL Press, Oxford, UK.

Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature,* **358**, 302–307.

Rischel, D., Thyberg, P., Rigler, R. & Poulsen, F. M. (1996). Time-resolved fluorescence studies of the molten globule state of apomyoglobin. *J. Mol. Biol.* **257**, 877–885.

Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energies determined by the linear extrapolation method, 1. Unfolding of phenylmethane sulfonyl α-chymotrypsin using different denaturants. *Biochemistry,* **27**, 8063–8068.

Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry,* **22**, 4690–4696.

Schmid, F. X. (1989). Spectral methods of characterizing protein conformation and conformational changes. In *Protein Structure: a Practical Approach* (Creighton, T. E., ed.), pp. 251–285, IRL Press, Oxford, UK.

Shastry, M. C. & Udgaonkar, J. B. (1995). The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* **247**, 1013–1027.

Silow, M. & Oliveberg, M. (1997). Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl Acad. Sci. USA,* **94**, 6084–6086.

Szabo, A. G. (1969). Kinetic characterization of complex reaction systems. In *Comprehensive Chemical Kinetics* (Bamford, C. H. & Tipper, C. F. H., eds), pp. 1–80, Elsevier, Amsterdam.

Tonomura, B., Nakatani, H., Ohnishi, M., Yamaguchi-Ito, J. & Hiromi, K. (1978). Test reaction for a stopped-flow apparatus. *Anal. Biochem.* **84**, 370–383.

Uversky, V. N. (1993). Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry,* **32**, 13288–13298.

Wildegger, G. & Kiefhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *J. Mol. Biol.* **270**, 294–304.

Zaidi, F. N., Nath, U. & Udgaonkar, J. B. (1997). Multiple intermediates and transition states during protein unfolding. *Nature Struct. Biol.* **4**, 1016–1024.

Edited by P. E. Wright

(Received 5 August 1997; received in revised form 12 November 1997; accepted 14 November 1997)