Published in final edited form as:
J Phys Chem B. 2012 June 14; 116(23): 7014–7025. doi:10.1021/jp3012365.

Microsecond Folding Dynamics of Apomyoglobin at Acidic pH

Ming Xu[†], **Olga Beresneva**[†], **Ryan Rosario**[†], and **Heinrich Roder**[†,‡,*]
[†]Fox Chase Cancer Center, Philadelphia, PA 19111

[‡]Department of Biochemistry and Biophysics, University of Pennsylvania, PA 19104

Abstract

Apomyolgobin (apoMb) is an important model for understanding the folding mechanism of helical proteins. This study focuses on a partially structured state of sperm whale apoMb populated at pH 4.2 (*M*-state), which structurally resembles a late kinetic intermediate in the formation of the native state (*N*) at higher pH. The thermodynamics and cooperativity of apoMb folding at pH 4.2 and 6.2 were studied by global analysis of the urea-induced unfolding transitions monitored by tryptophan fluorescence and circular dichroism. The kinetics of folding and unfolding of apoMb at pH 4.2 was measured over a time window from 40 to 850 μs, using fluorescence-detected continuous-flow measurements. Our observation of biphasic kinetics provides clear evidence for rapid (<100 μs) accumulation of previously unresolved intermediate states in both refolding and unfolding experiments. Quantitative kinetic modeling of the results, using a four-state mechanism with two intermediates on a direct route between the unfolded and folded states ($U \leftrightarrow I \leftrightarrow L \leftrightarrow M$), gave new insight into the conformational states and barriers that precede the rate-limiting step in the formation of the *N*-state of apoMb.

Keywords

protein folding; myoglobin; rapid mixing; continuous flow; fluorescence; circular dichroism

Introduction

There has been impressive progress in recent years in understanding the mechanisms of folding of small proteins, which often undergo highly concerted (two-state) folding transitions. Kinetic studies have identified several mini-proteins that fold within microseconds[1], making these complex reactions accessible to detailed computational analysis by all-atom molecular dynamics simulation[2–3]. However, despite decades of research[4–7], our understanding of larger proteins, which often exhibit complex, multi-stage folding reactions, remains incomplete. In particular, the role of partially structured states (intermediates) in folding remains a controversial issue[8]. Folding intermediates may serve productive roles during protein folding, for example by allowing independent folding of subdomains or acting as a hub for the vast number of microstates. Alternatively, premature structure formation may lead to kinetically trapped or misfolded states that can hinder productive folding or promote aggregation.

Progress in protein folding has been closely linked with technical advances in spectroscopy and fast kinetics. T-jump and other laser-initiated techniques[9–12] have provided rich kinetic information on the formation of α-helices and β-hairpins, which spans the high ns to low μs time range, as well as the folding of some small proteins and domains[13–17]. Advances in

*Corresponding author. roder@fccc.edu.

rapid mixing and detection methods prompted a renaissance of continuous-flow techniques[7,18–23]. By coupling an efficient capillary ball mixer with a sensitive CCD camera, we can routinely monitor fluorescence or absorbance changes with a dead time as short as 40 μs, which has yielded a wealth of information on early stages of protein folding[7,24–28].

Myoglobin is a heme protein consisting of eight helices (A through H). Its apo-form, apomyoglobin (apoMb), has long served as a paradigm for understanding the folding of helical proteins[29–30]. ApoMb adopts a compact structure very similar to that of the holoprotein, including an extensive hydrophobic core consisting of helices A, G and H, hydrophobic clusters in the region of helices B through E and native-like tertiary contacts among these helices[31–32]. Removal of the heme group, however, results in some structural perturbations, including loss of the F-helix, which provides the distal His-heme ligand of the holoprotein, as well as local disorder at the N-terminus of the G-helix and some loops. A non-native acidic form ("pH 4 intermediate", which we call "*M*-state") observed in equilibrium studies as a function of pH[33–35] resembles a kinetic intermediate encountered during folding at pH 6[36]. The observation of a burst phase (missing amplitude) in stopped-flow and quench-flow experiments[36–37] suggested that a second intermediate, *I*, accumulates on the sub-millisecond time scale, which was subsequently confirmed by ultrafast mixing[23,38–41]. The folding mechanism of apoMb can thus be described in terms of a sequential four-state scheme (Figure 1)[30,40,42]. Based on the protection of amide protons observed in their quenched-flow H/D exchange measurements, Nishimura at al.[43] concluded that the early folding event ($U \leftrightarrow I$) involves hydrophobically driven association of helices A, G and H. The B and E-helices contribute a few residues to the structural core, but remain otherwise unfolded at this stage. Stabilized by tertiary contacts with the AGH core, the B and E helices are fully formed in a subsequent folding step ($I \leftrightarrow M$, $\tau \sim 5$ ms). Time-resolved SAXS and CD measurements with a dead time of ~300 μs showed that the $U \leftrightarrow I$ transition is accompanied by a large decrease in chain dimensions and increase in helix content[40–41]. A further increase in the average helicity during the $I \leftrightarrow M$ transition is consistent with formation of the B and E helices and parts of C on the ms time scale[43]. The final stages of folding ($M \leftrightarrow$ N transition) involve further compaction and helix growth, as well as annealing of possible non-native contacts[44].

This study focuses on the *M*-state, the predominant equilibrium state of sperm whale apoMb at pH 4, which is thought to represent a stable analog to a kinetic intermediate encountered during formation of the native structure at higher pH[36,45]. Since the *M*-state maintains key features of the folded apoMb structure, including a compact globular fold with a native-like helical core, investigating its folding mechanism can thus provide clues about the early stages of apoMb folding. The rates of folding and unfolding at pH 4 are beyond the reach of conventional stopped-flow and quenched-flow techniques, but within the time resolution of our continuous-flow mixing apparatus. In addition to the rate-limiting steps in formation and unfolding of the *M*-state reported previously[38], we observed additional phases with time constants of 100 μs or less in both folding and unfolding experiments at pH 4.2, indicating that the process is more complex than previously thought. Using quantitative kinetic modeling, we show that the fluorescence-detected folding/unfolding kinetics as a function of urea concentration is accurately described by a four-state mechanism, $U \leftrightarrow I \leftrightarrow L \leftrightarrow M$, where *I* is an early intermediate in refolding and *L* is a previously undetected intermediate encountered during unfolding.

Materials and Methods

Materials

An expression plasmid for wild-type sperm whale apoMb was a gift from Drs. H. Jane Dyson and Peter Wright (The Scripps Research Institute). Ultra-pure urea was from MP Biochemicals, Inc. (Solon, OH). Trifluoroacetic acid was from Sigma-Aldrich (St. Louis, MO). Sodium acetate, acetonitrile, HPLC grade methanol and water, and other chemicals were from Fisher Scientific (Pittsburg, PA).

Expression and purification of apomyoglobin

Methods of Jennings et al[46] were followed with modifications. Wild-type sperm whale myoglobin gene cloned into pET17b (Novagen®) was expressed in *E. coli.* BL21 (DE3) following standard protocols. The cells were collected by centrifugation 4 hours after IPTG induction. The cell pellet was re-suspended in the lysis buffer (50 mM Tris, 10 mM EDTA, 100 mM NaCl, pH 8.0) and then lysed by repeated sonication. The insoluble portion was collected by centrifugation, washed three times with lysis buffer, and solubilized in 50% acetonitrile, 0.1% trifluoroacetic acid. The solution was clarified by centrifugation followed by passing through a 0.22-μm filter. The clarified solution was diluted with solution A (50% methanol, 0.1% trifluoroacetic acid) before being loaded onto an ACE preparative C4 reversed phase HPLC column (Advanced Chromatography Technologies, Aberdeen, Scotland). ApoMb was eluted with a gradient from 100% solution A to 67% solution B (100% methanol, 0.1% trifluoroacetic acid) monitored by absorbance at 280 nm. The purity was checked by SDS-PAGE followed by Coomassie Brilliant Blue staining. Purified protein was dialyzed extensively against de-ionized water, concentrated, flash froze in liquid nitrogen, and stored at −80 °C until use.

Folding/unfolding equilibrium

The urea-induced unfolding transition of the apoMb pH 4 intermediate (*M*-state) was monitored by tryptophan fluorescence and circular dichroism (CD) spectroscopy. ApoMb solutions of identical concentration in buffer without urea (native solution, 20 mM sodium acetate, pH 4.2) or with urea (unfolding solution, ~8 M urea, 20 mM sodium acetate, pH 4.2) were prepared. The concentration of urea was increased by manually withdrawing and adding aliquots of the unfolded solution into the native solution. The urea concentration was determined from the measured refractive index of withdrawn aliquots on a Reichert-Jung Mark II Abbe refractometer (Leica Inc., New York).

Tryptophan fluorescence was measured on a PTI QM-2000 spectofluorometer (Photon Technology International, Inc., South Brunswick, NJ) with a thermostated cuvette holder, using a 10 mm x 10 mm cuvette at 11 °C. Protein concentration was ~4 μM. Trp fluorescence emission spectra in the range of 300 ~ 450 nm were measured with an excitation wavelength of 288 nm. The bandwidths for excitation and emission were 1 nm and 4 nm, respectively. All data were averages of two scans. Background of matching solvent was subtracted before the spectra were normalized relative to the maximum intensity.

The urea-induced unfolding of apoMb was also measured by far-UV CD spectroscopy in the range of 225 – 250 nm using an AVIV 62DS spectropolarimeter (Lakewood, NJ). A 10-mm cuvette thermostated at 11 °C was used. Protein concentration was ~ 2 μM. The spectra were integrated and normalized against the integral value at 0 M urea.

Global analysis of equilibrium unfolding

The set of normalized fluorescence spectra at difference urea concentrations were transposed into urea-dependent fluorescence intensity at individual wavelengths. Using a global fitting procedure[47] implemented in IGOR Pro software (WaveMetrics Inc., Lake Oswego, OR), the family of the folding transitions at different wavelengths were fitted to a three-state model, $U \leftrightarrow I_{eq} \leftrightarrow M$, where U is the unfolded state, I_{eq} is an equilibrium intermediate at pH 4.2, and M is the predominant equilibrium state at pH 4.2. In the global fitting procedure, the thermodynamic parameters, C_{m1}, m_1, C_{m2}, and m_2 values, were treated as global parameters, i.e., one set of the C_m and m values were used to fit the family of the transitions in the transposed dataset. For each transition, C_{mi} is the denaturant concentration at which two inter-converting states are of equal population, and m_i is the slope of the free energy change as a function of denaturant concentration, which reports on the changes in solvent accessible surface area (SASA) accompanying structural transitions[48]. The C_m and m values were calculated based on the linear free energy relationship:

$$\Delta G_{ij}=\Delta G_{ij}(0)-m_{ij}\cdot C=m_{ij}(C_{m_{ij}}-C) \quad (1)$$

where ΔG_{ij} is the free energy of unfolding from *i*- to *j*-state, $\Delta G_{ij}(0)$ is the free energy of unfolding in the absence of urea, m_{ij} is the slope of the transition, C_{mij} is the mid point of the transition, and C is urea concentration. The intercepts and the slopes for the Trp fluorescence of the equilibrium states at individual wavelengths were treated as local parameters—they were allowed to vary for each urea-dependent transition. Consequently, by plotting the intercepts *vs.* wavelengths of each state, we generated fitted intrinsic Trp fluorescence spectra for each equilibrium state. The global three-state analysis was also conducted on a hybrid dataset in which the unfolding transition monitored by both fluorescence and far-UV CD were combined.

Kinetic measurements

We used a continuous-flow capillary mixing apparatus in fluorescence mode[20] to observe the sub-millisecond folding/unfolding kinetics of apoMb at pH 4. Efficient turbulent mixing is achieved by coupling a capillary mixer to an observation channel with a square cross section of 250 x 250 μm^2. The temperature was controlled by running the buffer delivery line through a heat exchanger connected to an external circulating water bath in addition to purging chilled nitrogen gas into the sample compartment. The detection system was recently upgraded with a back-illuminated, UV-enhanced charge-coupled device (CCD) camera with >50% quantum efficiency in the range of 220–820 nm (Alta U47-UV, Apogee Instruments Inc., Roseville, CA). The folding/unfolding kinetics of apoMb was monitored by intrinsic tryptophan fluorescence through a 300-nm blocking edge BrightLine® long-pass filter with >85% transmittance at 308–420 nm (Semrock, Inc., Rochester, NY). The 297-nm mercury line from a 300-W mercury-xenon arc lamp was used as the excitation source. Background subtraction and data averaging was performed using routines written in Matlab (MathWorks, Natick, MA).

The folding reaction was initiated by diluting acid-denatured protein solution (pH 2.0 in HCl) with refolding buffer (20 mM sodium acetate) with various urea concentrations to reach a final pH of 4.2. The unfolding reaction was started from apoMb in 20 mM sodium acetate, pH 4.2, diluted with the same buffer containing various concentrations of urea. The mixing ratio was 1:5 (protein solution:buffer) with a linear flow rate of 22.7 m/s in the observation channel. Final protein concentrations were ~10 μM. All experiments were conducted at 11 °C. For each combination of initial and final urea concentration, a set of matching solvent background were measured and subsequently subtracted. The dead time of

the mixer was 39 ± 4 μs. The time-dependent Trp fluorescence intensity was normalized to the intensity of the *U*-state (acid-unfolded apoMb at pH 2.0).

Kinetic modeling

A general matrix approach[7,49] was implemented with IGOR Pro software to evaluate the microscopic (elementary) rate constants and kinetic amplitudes. The microsecond kinetic folding/unfolding data are modeled using a four-state sequential model:

$$\mathrm{U} \underset{k_{\mathrm{IU}}}{\overset{k_{\mathrm{UI}}}{\rightleftharpoons}} \mathrm{I} \underset{k_{\mathrm{LI}}}{\overset{k_{\mathrm{IL}}}{\rightleftharpoons}} \mathrm{L} \underset{k_{\mathrm{ML}}}{\overset{k_{\mathrm{LM}}}{\rightleftharpoons}} \mathrm{M}$$

(Scheme 1)

where *U* is the unfolded state, *I* and *L* are kinetic intermediates, and *M* is the major folded state at pH 4.2. In order to calculate the microscopic rate constants defined in Scheme 1, a rate matrix is constructed based on a set of linear differential equations that describe the four-state kinetic model. The microscopic rate constants for the transition from *i*-state to *j*-state, k_{ij}, is defined by

$$\ln k_{ij} = \ln k_{ij}(0) + m^{\ddagger}_{ij} C/RT \quad (2)$$

where $k_{ij}(0)$ is the rate constant in the absence of urea, $m^{\ddagger}_{ij}$ is the kinetic m-value describing the urea-dependence of k_{ij}, C is the concentration of urea, R is the gas constant, and T is the absolute temperature. The observable rates (the macroscopic rate constants) correspond to the eigenvalues of the rate matrix, λ_i (*i*=1–3). The amplitudes associated with each eigenvalues are calculated from the eigenvectors of the rate matrix and the observable signal for each state. The calculation also yields equilibrium values that can be used to compare with the equilibrium urea unfolding curve for apoMb.

Free energy diagram

Free energy diagram was calculated using kinetic parameters derived from the kinetic modeling (Figure 8). The activation energy required to cross the barrier between the state *i* and *j* was calculated using the *M*-state as the reference state:

$$\Delta G^{\ddagger}_{ij} = -RT \ln(k_{ij}(0)/A_{\mathrm{o}}) - m^{\ddagger}_{ij} C \quad (3)$$

where A_{o} is the Arrhenius pre-exponential factor, which was assumed to be $5 \times 10^5 \mathrm{s}^{-1}$, based on estimates by Eaton and colleagues[1] (this choice has no effect on the conclusions). A cumulative sum of $m^{\ddagger}_{ij}$ in going from *M* to a given state *k* (transition state or intermediate), normalized by the sum of all $m^{\ddagger}_{ij}$ (corresponding to the equilibrium *m*-value) defines a parameter α_k, which is a measure of the change in solvent-accessible surface area for each state relative to *M*. In the case of a two-state transition, the α-value for the transition state corresponds to the Tanford β value[50].

Results

Equilibrium unfolding of apoMb

To determine the thermodynamic properties of the acidic form of apoMb (*M*-state), we measured the changes in tryptophan fluorescence and far-UV CD at pH 4.2 as a function of urea concentration. The fluorescence emission spectra of apoMb at each urea concentration

were measured at 11°C (Figure 2A). The protein contains two tryptophan residues, Trp7 and Trp14, both in helix A. Trp14 is a conserved residue that participates in forming the early hydrophobic core structure[51] onto which tertiary contacts promoting further helix formation are formed. Both tryptophans are buried in the hydrophobic interior of the folded protein, and exposure of the tryptophans to the solvent upon unfolding results in a red-shifted emission spectrum and reduced fluorescence intensity (Figure 2A). The fluorescence spectra gain intensity with increasing urea concentrations up to ~1 *M*, and then decrease continuously with further addition of urea (Figures 2A and 2B), suggesting the presence of a hyperfluorescent intermediate.

In order to globally fit the urea unfolding data to a thermodynamic model, the set of urea-dependent Trp fluorescence spectra was transposed, yielding a series of unfolding transitions at individual wavelengths (Figure 2B). Although most individual transitions can be fitted to a two-state model, the thermodynamic parameters obtained vary significantly from wavelength to wavelength, which is a clear indication that at least three states are populated at different urea concentrations. The data were therefore fitted to a three-state equilibrium model ($U \leftrightarrow I_{eq} \leftrightarrow M$), adopting global fitting procedures developed by Latypov et al.[47]. The thermodynamic parameters (C_m and m values for each transition) are treated as global fitting parameters. The optimized set of parameters (Table 1) defines the evolution of each equilibrium state as a function of the urea concentration (Figure 2D). A non-native equilibrium state, I_{eq}, is well populated even in the absence of urea, confirming prior findings[37] that apoMb is structurally heterogeneous at pH 4.2 (our *M* and I_{eq} states were previously called Ia and Ib, respectively). The population of I_{eq} peaks at ~ 1 M urea while that of *M* decreases continuously, and *U* becomes the predominant species at urea concentration higher than 3.5 *M*.

In our global fitting procedure, the y-axis intercepts and slopes vs. urea concentration of the intrinsic fluorescence or CD signal of each state are treated as local fitting parameters. If we plot the intercepts (i.e., the fluorescence signal extrapolated to [urea] = 0) for a given state (*M*, I_{eq} or *U*) as a function of wavelength, we obtain the intrinsic fluorescence spectrum of that state in the absence of denaturant (Figure 2C). The intrinsic spectra of *M* and I_{eq} cannot be measured directly because of the heterogeneous nature of apoMb at pH 4.2. The fitted Trp fluorescence spectrum of I_{eq} is more intense compared to that of *M*, confirming that the intermediate is hyperfluorescent, probably due to the loss of a specific contact with an intramolecular quencher. Along with the increasing population of I_{eq} at low urea concentration, its hyperfluorescence accounts for the initial increase in fluorescence intensity (Figure 2B). The maximum of the emission spectrum of I_{eq} is red-shifted by 7 nm relative to *M* (Figure 1C), indicating that Trp7 and Tpr14 are in a more solvent-exposed environment. The fact that the fluorescence properties of I_{eq} differ markedly from those of *M* and *U* makes the fitting procedure quite robust (nearly identical solutions are obtained for different initial parameter sets). Due to the marginal stability of apoMb at pH 4.2 and simultaneous population of the two states, there is no native baseline region, and we assumed that the slopes for both *M* and I_{eq} vs. urea concentration are the same as those of the *U*-state observed at urea concentration above 3.5 *M* (Figure 2B). This assumption is supported by the equilibrium transitions for apoMb at pH 6.2, which show similar baseline slopes both at low and high urea concentrations (see Figure 4).

The urea-induced apoMb unfolding transition was also monitored via the changes in the far-UV CD spectrum. To obtain a more precise measure of the relative helix content, we integrated the CD spectrum between 225 and 250 nm (red circles in Figure 3B) and normalized the data, using the data at pH 6.2 to estimate the native baseline slope (note that lower wavelength are masked by the absorbance of urea). In Figure 3A, we compare the results with fluorescence-detected unfolding curves at two characteristic wavelengths. The

normalized signal at 315 nm, where the fluorescence spectra of *M*- and I_{eq} cross each other (Figure 2C), measures the combined population of *M* and I_{eq} relative to *U*, whereas the signal at 365 nm (corresponding to the intercept between the *M*- and *U*-state spectra) is dominated by the population of I_{eq}. At urea concentrations between 0 and 1.5 M, we observe a sharp decrease in the CD signal consistent with a ~40% loss in helix content as the population shifts from *M* towards I_{eq}. However, at urea concentrations above 1.5 *M*, the CD-detected helix content is systematically higher than the population of structured states ([*M*] + [I_{eq}]) monitored via fluorescence. This behavior is inconsistent with a three-state mechanism, and we had to introduce an additional equilibrium intermediate to account for the CD data. Indeed, a four-state unfolding mechanism ($M \leftrightarrow L_{eq} \leftrightarrow I_{eq} \leftrightarrow U$) accurately represents the combined fluorescence and integrated CD data (Figure 3), as well as the CD-detected transitions at each wavelength (data not shown). Global optimization of equilibrium parameters (Table 1) leads to a plausible distribution of the states vs. urea concentration (panel D), and the local fitting parameters predict distinct intrinsic fluorescence spectra for all four states (Figure 3C).

To determine the stability and cooperativity of apoMb under more stabilizing conditions, we recorded a full set of equilibrium data, including fluorescence emission and CD spectra vs. urea concentration, at pH 6.2. The fluorescence-detected unfolding curves again exhibit wavelength-dependent behavior with two distinct transitions (Figure 4A). In contrast to the behavior at pH 4.2, the CD-detected unfolding transition coincides with those detected by fluorescence, and the combined data set can be modeled accurately using a three-state equilibrium model ($N \leftrightarrow I_{eq} \leftrightarrow U$). The fluorescence spectra obtained by the global analysis (Figure 4B) show that the equilibrium intermediate at pH 6.2 is a highly fluorescent state with a ~2-fold higher yield than the N-state and a peak wavelength of 330 nm, intermediate between those of the N-state (323 nm) and the *U*-state (346 nm). I_{eq} accumulates to detectable levels between 2 and 5 M urea and reaches a maximum population of about 20% at 3.5 M urea (Figure 4C).

Microsecond folding/unfolding kinetics of apoMb

Despite numerous kinetic studies[36,39–40,52–53], we do not have a complete picture of the kinetic folding mechanism of apoMb. Indirect evidence for accumulation of intermediate(s) on the sub-millisecond time scale has previously been obtained on the basis of missing amplitude (burst-phase) observations in fluorescence-detected stopped-flow[38], continuous-flow circular dichroism and SAXS[40], and quench-flow hydrogen exchange[41] experiments. In order to directly observe these putative fast events in folding, we used a continuous-flow mixing apparatus with a ~40 μs dead time[20] to measure the changes in the intrinsic Trp fluorescence associated with folding and unfolding of apoMb at pH 4.2 as a function of urea concentration (Figure 5). The folding traces at low urea concentrations (0 – 2.14 *M*) are accurately represented as a sum of two exponentials (Figure 5A, solid lines). The biphasic nature of the folding kinetics at pH 4.2 signifies the involvement of at least one folding intermediate. Surprisingly, the unfolding reaction of apoMb also proceeds in a biphasic fashion with a fast increase in fluorescence (~10,000 s^{-1}) followed by an exponential decay with a rate of ~3,000 s^{-1} (Figure 5B). This unusual feature is clearly not due to instrumental artifacts because: 1) a matching solvent background was measured and subtracted from each trace; 2) scattering artifacts due to incomplete mixing (schlieren effects) were previously shown to decay with an apparent time constant of ~10 μs[28], while we observe here is an exponential growth with a time constant in the range of 30 ~ 140 μs, which falls within the time resolution of our apparatus (~40 μs dead time). Unfolding intermediates are uncommon, although some examples have been reported previously[28,38,54–55].

Figure 6A shows a plot of the logarithm of the apparent folding/unfolding rate constant as a function of urea concentration (chevron plot). Three phases with urea-dependent rates in the

sub-ms time range are observed. The fastest phase detected in the folding experiments represents mainly the transition between U and I (red circles, Figure 6). The rate profile for the slowest phase (blue circles and squares, Figure 6) describes a V-shape with a curved unfolding branch (roll-over), which can be attributed to an intermediate on the unfolding pathway[54–56]. In these previous reports, a single phase accounts for the total change in signal upon unfolding, indicating that the intermediate responsible for the roll-over is a poorly populated high-energy state. In contrast, the biphasic unfolding traces we observed for apoMb (Figure 6) indicate transient accumulation of a kinetic intermediate, i.e. a state that is comparable to or lower in free energy than the folded state.

Kinetic modeling

The improved time resolution of our continuous-flow apparatus allowed us to monitor the folding/unfolding kinetics of the acidic form of apoMb in the microsecond time regime. In addition to the main folding/unfolding phase on the sub-ms time scale (λ_3), we directly observed two distinct kinetic phases with rate constants well over 10,000 s^{-1} during folding and unfolding, respectively (λ_1 and λ_2, Figure 6A). Since a minimum of four states are need to yield three observable rate constants, we used a sequential four-state mechanism (Scheme 1) to model the folding/unfolding kinetics of apoMb at pH 4.2 (Figure 6), using the rate-matrix approach described in *Material and Methods*.

The first aim of our kinetic modeling effort was to optimize the microscopic rate constants for each transition in Scheme 1, k_{ij}, and their urea-dependence, expressed in terms of kinetic m values, $m^{\ddagger}_{ij}$ (dashed lines) in order to reproduce the observed rate constants and their dependence on urea concentration (Figure 6A). General solutions of a four-state kinetic mechanism can be readily obtained numerically. However, implementing least-squares fitting is difficult because certain constellations of input parameters lead to degenerate solutions, and we chose to manually adjust the kinetic parameters, k_{ij} and $m^{\ddagger}_{ij}$ in order to fit the observed rate constants. We then assigned initial fluorescence yields and urea-dependent slopes for each of the four states in Scheme 1 and varied them in an effort to reproduce the observed amplitudes during folding (Figure 6B) and unfolding (Figure 6C). As a more stringent test of our model, we used this preliminary set of kinetic and spectroscopic parameters to globally "fit" the family of folding and unfolding traces (Figure 7). A constrained least-squares fitting procedure allowed us to make final adjustments in the intrinsic fluorescence parameters of each state. After several rounds of manual and least-squares optimization, we were able to find a common set of elementary rate constants and kinetic m-values (Table 2), as well as consistent sets of spectroscopic parameters (Table 3) that qualitatively reproduced the observed time course of folding and unfolding (two sets of spectroscopic parameters were necessary to fit folding and unfolding data, respectively, because of differences in conditions and normalization errors). This global fit (lines in Figure 7) predicts all salient features of the kinetics, including the biphasic increase in fluorescence during folding and the bell-shaped time course of unfolding. A free energy diagram constructed on the basis of the optimized kinetic parameters (Table 2) is shown in Figure 8.

In an effort to assess the uniqueness of our kinetic model (Scheme 1) and parameter set (Table 2), we explored several alternative solutions and kinetic schemes. For example, we identified a distinct set of elementary rate parameters that also reproduces the urea-dependence of observable rates (Figure 6A) where k_{ML} rather than k_{LI} is the rate-limiting step in unfolding. However, this solution failed to predict a key feature of the observed unfolding kinetics, namely the reversal in the sign of the amplitude factors for the two predominant unfolding phases, A_2 and A_3 (Figure 6B), and thus was unable to reproduce the observed time course of unfolding (Figure 7B). We also tested some alternative four-state mechanisms, including a T-shaped schemes with an off-pathway state and a diamond-

shaped scheme with two parallel paths connecting U to M, none of which yielded a satisfactory fit of the observed rates and amplitudes. While an exhaustive analysis of all possible combinations was not done, we are confident that Scheme 1 is the correct mechanism. In particular, an obligatory intermediate preceding the rate-limiting unfolding barrier, *L*, is necessary to reproduce the biphasic unfolding behavior. Similarly, the double-exponential folding traces with a fast phase λ_1, which is distinct from the fast unfolding phase, λ_2, provide strong evidence for the presence of an obligatory folding intermediate on a direct path between *U* and *M*.

Discussion

Structural heterogeneity of apoMb at pH 4.2

Griko et al.[33–34] showed that acidification of apoMb gives rise to a two-step unfolding transition with a prominent intermediate state populated near pH 4. Subsequent mutational and H/D exchange NMR studies showed that this “pH 4 intermediate” contains a compact hydrophobic core comprising the A, G, and H helices while the rest of the molecule appears largely unfolded[35,57], and structurally resembles a kinetic intermediate observed during folding of apoMb at pH 6[36]. Following earlier evidence for conformational heterogeneity[58], Jamin et al.[37–38] showed that apoMb populates two distinct equilibrium states at pH 4.

When comparing our global three-state analysis of apoMb at pH 4.2 with that at pH 6.2 (Table 1), it is apparent that the unfolding free energy of the *M*-state (2.2 kcal/mol) is much lower than that of *N* (8.4 kcal/mol). Possible factors contributing to the low structural stability of *M* compared to *N* are unfavorable electrostatic interactions due to the increase in net charge upon lowering the pH from 6 to 4, together with the fact that the *M*-state contains only a subset of the native helices and tertiary interactions. Our thermodynamic data further show that the free energy of the equilibrium intermediate at pH 6.2 (2.9 kcal/mol) is closer to *U* than N and comparable to that of the *M*-state at pH 4.2, consistent with a previous report by Barrick and Baldwin[34].

Our three-state equilibrium analysis of the fluorescence data (Figure 2) confirms and extends prior findings[37] that at moderately acidic pH apoMb can adopt two distinct folded states, *M* and I_{eq}, whose populations are highly sensitive to solution conditions, including pH, denaturant and ionic strength. In the absence of denaturant, *M* and I_{eq} account for 60% and 40% of the total population, respectively (Figure 2C). The population of I_{eq} increases with increasing urea concentration at the expense of *M* and reaches a maximum (60%) at ~1 M urea, at which point *M* and *U* have ~20% population each. The fact that neither *M* nor I_{eq} are fully populated under any condition makes it impossible to directly measure their spectral properties. However, we can mathematically extract the spectra of each state by global analysis of the urea-unfolding data, which yields deconvoluted fluorescence emission spectra in the absence of denaturant (Figure 2D). The fluorescence spectrum of *M* is more intense and blue-shifted relative to that of *U* (326 nm vs. 345 nm), indicating that the two tryptophans in the myoglobin sequence, Trp7 and Trp14, are largely inaccessible to the solvent in the *M*-state. The emission maximum of I_{eq} (332 nm) is intermediate between that of *M* and *U* (Figure 2D), suggesting that the tryptophans are partially solvent-exposed. However, the fact that the peak intensity is substantially lower in *M* compared to I_{eq} can be explained by the presence of an intramolecular quenching interaction, which is relieved in I_{eq}. Given the close proximity of Trp7 and Trp14 in the folded structure of Mb (~9.5 Å), a likely contribution to quenching is Trp-Trp energy transfer (self-quenching), in addition to other known intramolecular quenchers, such as His, Lys and Glu[59]. Jamin et al suggested that both Trp residues are partially buried in the pH 4 intermediate, based on the effect of Trp → Phe mutations[38]. On the other hand, Haruta and Kitagawa[60] concluded that Trp7 of horse apoMb is fully exposed in the pH 4 intermediate while Trp14 remains buried, based

on time-resolved UV resonance Raman (UVRR) spectra, which indicate that both Trp residues are transiently buried in a hydrophobic microenvironment prior to formation of *M* while Trp7 becomes exposed once *M* is formed. Although horse and sperm whale apoMb share 88% sequence identity and show similar folding behavior, subtle differences in local structure and stability may give rise to different fluorescence properties.

Although the thermodynamic stability of *M* is much lower than that of N (Table 1), the two states share some structural similarities: a) 44% of the residues are helical in *M*, compared to 53% in N[40]; b) over 75% of the solvent-accessible surface area is buried in *M* (m_{UM} = 1.93 kcal/mol/*M*, compare to m_{UN} = 2.53 kcal/mol/*M*, Table 1). On the other hand, when comparing the intrinsic fluorescence spectra of the equilibrium states at pH 4.2 and 6.2, using the *U*-state as a reference state (Figure 9), we note that the acidic *M*-state has a significantly higher relative fluorescence yield (1.72) compared to N (1.25) and a red-shifted emission maximum (326 vs. 323 nm). This suggests that *M* is a more loosely packed state with partially relaxed Trp-quencher interactions. This loss of specific tertiary packing interactions is one of the characteristics of a molten globule state[61]. At the same time, the emission spectra of the predominant equilibrium intermediate derived from the three-state analyses at pH 4.2 and pH 6.2 are remarkably similar (Figure 9), indicating that a common intermediate with similar structural properties accumulates under both conditions.

The acidic and neutral forms of apoMb show marked differences in terms of the cooperativity of their unfolding transitions. At pH 6.2, a three-state equilibrium ($N \leftrightarrow I_{eq} \leftrightarrow U$) fully accounts for both the fluorescence- and CD-detected unfolding transitions, and the population of I_{eq} remains below 20% throughout the transition (Figure 4). In contrast, the deviations from two-state equilibrium behavior are much more pronounced at pH 4.2, and the three-state equilibrium analysis predicts an intermediate accounting for as much as 60% of the population (Figures 2 and 3). However, the CD data are incompatible with the three-state model fitted to the fluorescence data (Figure 3A), and a fourth state has to be introduced to account for the combined data (Figure 3B). All four states have distinct fluorescence emission spectra with *L* having the highest yield (Figure 3C). A constrained fit of the normalized CD data in Figure 3A using the global fitting parameters (Table 1) yields relative helix contents of 1, 0.79, 0.62 and 0 for *M*, *L*, *I* and *U*, respectively, indicating a progressive loss of native secondary structure with each unfolding step.

Kinetic mechanism of folding of apoMb

Thermodynamics alone is insufficient for fully characterizing a folding mechanism. In order to determine the connectivity of states and the height of free energy barriers (transition state ensembles), we carried out a detailed analysis of the kinetics of folding and unfolding of apoMb as a function of urea concentration. By exploring the kinetics of formation and unfolding of the *M*-state at pH 4.2, which is thought to represent a stable analog of a late intermediate in the formation of the native apoMb structure[36,45,62], we can gain insight into early steps in folding that may be invisible at higher pH, where folding is a more cooperative process. This, along with the fact that the *M*-state features key attributes of a folded state in terms of its compactness, high secondary structure content and critical tertiary (helix-helix) interactions, makes the findings relevant for understanding folding under more physiological conditions.

On the basis of fluorescence and CD measurements with a time resolution of ~400 μs, Weisbuch et al.[39] concluded that folding of apoMb at pH 4.2 occurs in two stages each of which can been described as cooperative processes. However, these stopped-flow experiments, as well as earlier continuous-flow fluorescence measurements with a 200 μs dead time[38,52], showed evidence for unresolved fluorescence changes, especially in

unfolding experiments, suggesting that additional intermediate(s) accumulate within the instrumental dead time. We were able to fully resolve the folding/unfolding kinetics of apoMb at pH 4.2 on our continuous-flow fluorescence instrument, which uses a capillary ball mixer to reach dead times as short as 40 μs. The kinetic traces measured during folding are clearly biphasic (Figure 5A), indicating that an intermediate is encountered within 30–40 μs of refolding. In addition, we obtained direct kinetic evidence for the transient accumulation of an unfolding intermediate, which gives rise to bell-shaped unfolding traces with a fast rise in fluorescence followed by a slower decay (Figure 5B).

We showed by quantitative kinetic modeling that the folding mechanism of apoMb at pH 4.2 can be described by a four-state scheme, $U \leftrightarrow I \leftrightarrow L \leftrightarrow M$ (Scheme 1), where *I* and *L* are obligatory intermediates on a direct path between *U* and *M*. The *I*-state is detectable only during folding while *L* is manifest in unfolding experiments. Since the three observable rate constants predicted by the model are relatively close to each other (within a factor of 10–20 at low to intermediate urea concentrations), the different phases are strongly coupled (Figure 6). The kinetic coupling is especially strong under folding conditions, and each observable process (λ_1–λ_3) is a complex function of several elementary rate constants. For example, four elementary rate constants, k_{UI}, k_{IU}, k_{IL} and k_{LM}, contribute to the main phase in refolding, λ_3. Therefore, an earlier model[37–39] in which the sub-millisecond folding kinetics of apoMb is attributed to the transition between the unfolded state and an early folding intermediate ($U \leftrightarrow I_a$ in the mechanism of Jamin et al.) has to be revised in light of our more complete kinetic data and quantitative kinetic analysis.

To gain further insight into the underlying mechanism, we used the optimized elementary rate constants and kinetic m-values to construct free energy diagrams at different denaturant concentrations (Figure 8). At low urea concentrations, *U* and *I* are comparable in free energy, and their interconversion via *TS1* dominates the fast folding phase ($\lambda_1 \cong k_{UI} + k_{IU}$). Although the model predicts three phases, the refolding traces cannot be fitted reliably to more than two exponentials. The second phase predicted by the model, λ_2, has low amplitude under refolding conditions (Figure 6B) and cannot be isolated. To overcome the limitations of empirical fitting of multi-exponential functions, we numerically solved the kinetic equations representing Scheme 1 to predict the time course of folding and unfolding curves at different urea concentrations (Figure 7). The fact that we were able to find a global solution, i.e., a single set of kinetic parameters that reproduces the whole family of folding and unfolding traces provides striking confirmation for the validity of the kinetic model.

Figure 8 shows that *M* has the lowest free energy in the absence of urea, and is thus the dominant equilibrium state. However, *L* is only slightly (~1 kcal/mol) higher in free energy, consistent with our observation that both states are populated at equilibrium (Figure 3D). At low urea concentrations, the late barriers in folding, *TS2* and *TS3*, are comparable in height and slightly smaller than the first barrier, *TS1*. As a result, all three elementary folding steps contribute to the net rate of folding, λ_3, with *TS2* and *TS3* making the dominant contributions. Because addition of urea preferentially stabilizes the more unfolded states, *U* and *TS1*, relative to the more compact states, crossing of *TS2* (k_{IL}) begins to dominate the rate of folding at urea concentrations of ~0.5 *M* and above. At low urea concentration the *I*-state is marginally more stable than *U* and accumulates as a transient intermediate with a rate constant $\lambda_1 \cong k_{UI} + k_{IU}$. As we jump to urea concentrations above the midpoint of the equilibrium transition (>1.5 M urea), the situation is reversed with the *L*-state acting as a transient intermediate in unfolding. Initially, *TS3* and *TS2* are similar in height, and both k_{ML} and k_{LI} contribute to the slow unfolding phase (λ_3). With increasing urea concentration, *TS2* becomes the dominant barrier in unfolding (λ_3 approaches k_{LI}) while *TS3* decreases in height, giving rise to an accelerating fast unfolding phase (λ_2 in Figure 6).

Kinetic m-values provide information on the differences in solvent accessible surface area (SASA) between the various states populated in folding/unfolding reactions, including transition-state ensembles[63]. Each state in Figure 8 is labeled with its α value defined in *Methods*, which reports on the relative changes in SASA ranging from that of the most structured state, *M* (α = 0) the fully unfolded state, *U* (α = 1). The transition from *U* to *I* is accompanied by a significant decrease in solvent exposure (32%), and formation of *L* results in an additional 34% reduction. Thus, the *L*-state is about 2/3 along the way towards *M* in terms of SASA. Figure 8 also sheds light onto the origin of the unusual biphasic time course of unfolding we observed. In several previous studies[54–56], a rollover in the unfolding arm of the chevron plot has been attributed to a high-energy intermediate combined with a change in rate-limiting barriers from a later to an earlier step in unfolding. If this were the case for apoMb, *TS3* would become the dominant barrier at high urea concentrations, resulting in single-exponential kinetics with a rate approaching k_{LM} under strongly denaturing conditions. In contrast, Figure 6 indicates that *TS3* is always smaller than *TS2*, and the rate of unfolding is limited by the rate of crossing *TS2* (k_{LI}). As a result, the *L*-state is not a high-energy intermediate, but accumulates transiently during unfolding, giving rise to biphasic unfolding traces (Figure 5B). The observation that the fluorescence traces increase at first before decaying at longer times indicates that the fluorescence yield of the *L*-state is higher than both *M* and *U* (Table 3), which is consistent with the 4-state analysis of the equilibrium data (Figure 3C).

According to Figure 8, *M, L* and *I* are similar in free energy at intermediate urea concentrations, and all three states are thus populated at equilibrium. The kinetic parameters predict that the *L*-state is the predominant equilibrium intermediate with a maximum population near 1 M urea, and that the *I*-state population peaks around 2 *M* (Figure 6D). This prediction is in qualitative agreement with the distribution of equilibrium states derived from the equilibrium unfolding data (Figure 3C), although the peak populations are substantially lower, mainly due to a shift in the population of the *M*-state toward higher urea concentration. A likely factor contributing to this discrepancy is cis/trans isomerization of proline peptide bonds. If some of the proline peptide bonds are restricted to their *trans* isomer in *M*, but can isomerizes in the more loosely folded *L* and *I* states at equilibrium, this will entropically stabilize *U*, *I* and *L* relative to *M* in equilibrium experiments. In contrast, the Pro peptide bonds maintain their native isomers in kinetic unfolding experiments, and the lack of entropic stabilization may thus account for the low population of kinetic intermediates.

The total free energy of unfolding calculated from the kinetic parameters in Table 2 (2.6 kcal/mol) is comparable to that derived from the three-state fit of the equilibrium data (2.2 kcal/mol, Table 1). However, the four-state equilibrium analysis yields a substantially higher value (4.9 kcal/mol). This discrepancy appears to be largely due to uncertainties in m-values; the kinetic analysis may underestimate m-values while the equilibrium analysis tends to overestimate m-values for poorly resolved transitions. Although we found no evidence for a burst-phase (missing amplitude) effect in our kinetic data, we cannot rule out the possibility that additional states accumulate within the instrumental dead time, which may account for some of the missing free energy.

Conclusions

Prior kinetic studies using optical probes or H/D exchange to follow folding of apoMb have identified an early intermediate during folding at pH 6 that resembles the predominant equilibrium state, *M*, at pH 4.2[36,38,62]. By focusing on the structural events involved in the formation of the *M*-state, we were able to gain more detailed insight into the early structural events that lead to the rate-limiting barrier in the formation of the *N*-state. The premise that the folding mechanism at acidic pH corresponds to a truncated version of the mechanism at

pH 6 is supported by our global analysis of the equilibrium unfolding data; while the *M*-state is distinct from *N* in terms of fluorescence properties and helix content, equilibrium intermediates with similar fluorescence spectra accumulate both at pH 4.2 and 6.2 (Figures 2–4 and 9). At pH 4.2 we detected two well populated (40–70%) equilibrium intermediates with distinct fluorescence and CD spectra, whereas only one marginally populated (<20%) intermediate was observed at pH 6.2. Thus, acidification leads to a loss in structural cooperativity, revealing partially structured states that cannot be resolved from the *N* and *U*-states that dominate the folding equilibrium at higher pH.

Previous studies hinted at the presence of early intermediates during folding and unfolding of apoMb at pH 4, but lacked the time resolution to directly monitor their formation[38–40]. By using a turbulent-mixing continuous-flow technique with Trp fluorescence detection, covering the time window from 40 to 850 μs, we were able to fully characterize the kinetics of structure formation and unfolding of apoMb at pH 4.2. Our observation of biphasic kinetics not only during refolding (<2 M urea), but also during unfolding (>2 M urea) indicates that two distinct kinetic intermediates accumulate during folding and unfolding, respectively. An intermediate, *I*, accumulates early in folding (τ = 30 - 50 μs), prior to the rate-limiting step in formation of the *M*-state with a urea-dependent time constant of 0.25 – 1 ms, and a second intermediate, *L*, accumulates during unfolding, giving rise to a transient increase in fluorescence on the 100 μs time scale followed by a slower decay (τ = 300 – 600 μs). All kinetic data, including the observable rate constants and kinetic amplitudes vs. urea concentration (Figure 6), as well as the combined set of folding/unfolding traces (Figure 7), can be modeled quantitatively using a sequential four-state mechanism (Scheme 1). It is interesting to note that the kinetic intermediates *I* and *L* also accumulate under equilibrium conditions (Figure 6D). This allowed us to study their structural properties in more detail by using steady state fluorescence or far-UV CD spectroscopy. In particular, the far-UV CD data in Figure 3 indicates that the folding transition from *U* to *M* via *I* and *L* is accompanied by a stepwise increase in α-helix content. At the same time, the fluorescence emission maximum moves to lower wavelengths with each stage of folding, indicating that Trp7 and Trp14 become increasingly more buried (Figure 3C). The observed increase in fluorescence yield for the first two folding steps is consistent with progressive compaction, whereas the final decrease upon formation of *M* suggests the formation of specific tertiary interactions.

Because of the strong coupling between the different kinetic phases, most elementary rate constants in Scheme 1 are well constrained by the data (Figure 6). Our inability to reproduce key features of the observed kinetics with alternative kinetic schemes strongly suggests that both *I* and *L* are obligatory intermediates on a direct path between *U* and *M*. This is a surprising finding, considering the large number of states and complex pathways seen in full-atom MD simulations of fast-folding proteins[2]. However, analysis of these simulations using Markov state model and other clustering algorithms has suggested that a limited number of states may serve as kinetic hubs, thus reducing the number of pathways[64]. More recently, based on their extensive MD simulations on 12 proteins ranging up to 78 residues in size, Lindorff-Larsen et al.[3] concluded that some proteins follow a single dominant folding route.

Acknowledgments

We thank Drs. H Jane Dyson and Peter Wright for providing an expression plasmid for sperm whale apomyoglobin and Dr. Chiaki Nishimura for his help with apoMb purification. We also thank Dr. Valentina Bychkova for providing a sample of apomyoglobin for initial continuous-flow experiments. The work was supported by grant MCB-0744607 from the National Science Foundation (to HR), and grant CA06927 and an Appropriation from the Commonwealth of Pennsylvania to the Fox Chase Cancer Center, and the Spectroscopy Support, Biotechnology, DNA Sequencing facilities.

References

1. Kubelka J, Hofrichter J, Eaton WA. Curr Opin Struct Biol. 2004; 14:76–88. [PubMed: 15102453]
2. Bowman GR, Voelz VA, Pande VS. Curr Opin Struct Biol. 2011; 21:4–11. [PubMed: 21081274]
3. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. Science. 2011; 334:517–520. [PubMed: 22034434]
4. Kim PS, Baldwin RL. Annu Rev Biochem. 1990; 59:631–660. [PubMed: 2197986]
5. Roder H, Colón W. Curr Opin Struct Biol. 1997; 7:15–28. [PubMed: 9032062]
6. Bilsel O, Matthews CR. Adv Protein Chem. 2000; 53:153–207. [PubMed: 10751945]
7. Roder H, Maki K, Cheng H. Chem Rev. 2006; 106:1836–1861. [PubMed: 16683757]
8. Sosnick TR, Barrick D. Curr Opin Struct Biol. 2011; 21:12–24. [PubMed: 21144739]
9. Huang CY, He S, DeGrado WF, McCafferty DG, Gai F. J Am Chem Soc. 2002; 124:12674–12675. [PubMed: 12392410]
10. Callender RH, Dyer RB, Gilmanshin R, Woodruff WH. Annu Rev Phys Chem. 1998; 49:173–202. [PubMed: 9933907]
11. Gruebele M. Annu Rev Phys Chem. 1999; 50:485–516. [PubMed: 15012420]
12. Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, Hofrichter J. Annu Rev Biophys Biomol Struct. 2000; 29:327–359. [PubMed: 10940252]
13. Wang M, Tang Y, Sato S, Vugmeyster L, McKnight CJ, Raleigh DP. J Am Chem Soc. 2003; 125:6032–6033. [PubMed: 12785814]
14. Zhu Y, Alonso DO, Maki K, Huang CY, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F. Proc Natl Acad Sci U S A. 2003; 100:15486–15491. [PubMed: 14671331]
15. Wang T, Zhu Y, Gai F. J Phys Chem. 2004; B108:3694–3697.
16. Bunagan MR, Yang X, Saven JG, Gai F. J Phys Chem B. 2006; 110:3759–3763. [PubMed: 16494434]
17. Buchner GS, Murphy RD, Buchete NV, Kubelka J. Biochim Biophys Acta. 2011; 1814:1001–1020. [PubMed: 20883829]
18. Chan CK, Hu Y, Takahashi S, Rousseau DL, Eaton WA, Hofrichter J. Proc Natl Acad Sci U S A. 1997; 94:1779–1784. [PubMed: 9050855]
19. Takahashi S, Yeh SR, Das TK, Chan CK, Gottfried DS, Rousseau DL. Nat Struct Biol. 1997; 4:44–50. [PubMed: 8989323]
20. Shastry MCR, Luck SD, Roder H. Biophys J. 1998; 74:2714–2721. [PubMed: 9591695]
21. Knight JB, Vishwanath A, Brody JP, Austin RH. Phys Rev Lett. 1998; 80:3863–3866.
22. Bilsel O, Kayatekin C, Wallace LA, Matthews CR. Rev Sci Instr. 2005; 76:014302.
23. Lapidus LJ, Yao S, McGarrity KS, Hertzog DE, Tubman E, Bakajin O. Biophys J. 2007; 93:218–224. [PubMed: 17416618]
24. Shastry MCR, Roder H. Nat Struct Biol. 1998; 5:385–392. [PubMed: 9587001]
25. Welker E, Maki K, Shastry MC, Juminaga D, Bhat R, Scheraga HA, Roder H. Proc Natl Acad Sci U S A. 2004; 101:17681–17686. [PubMed: 15574490]
26. Apetri AC, Maki K, Roder H, Surewicz WK. J Am Chem Soc. 2006; 128:11673–11678. [PubMed: 16939293]
27. Tzul FO, Kurchan E, Roder H, Bowler BE. Biochemistry. 2009; 48:481–491. [PubMed: 19113858]
28. Chen KC, Xu M, Wedemeyer WJ, Roder H. Biophys J. 2011; 101:1221–1230. [PubMed: 21889460]
29. Dyson HJ, Wright PE. Chem Rev. 2004; 104:3607–3622. [PubMed: 15303830]
30. Jamin M. Protein Pept Lett. 2005; 12:229–234. [PubMed: 15777270]
31. Eliezer D, Wright PE. J Mol Biol. 1996; 263:531–538. [PubMed: 8918936]
32. Lecomte JT, Sukits SF, Bhattacharya S, Falzone CJ. Protein Sci. 1999; 8:1484–1491. [PubMed: 10422837]
33. Griko YV, Privalov PL, Venyaminov SY, Kutyshenko VP. J Mol Biol. 1988; 202:127–138. [PubMed: 3172208]
34. Barrick D, Baldwin RL. Biochemistry. 1993; 32:3790–3796. [PubMed: 8466917]

35. Hughson FM, Wright PE, Baldwin RL. Science. 1990; 249:1544–1548. [PubMed: 2218495]
36. Jennings PA, Wright PE. Science. 1993; 262:892–895. [PubMed: 8235610]
37. Jamin M, Baldwin RL. J Mol Biol. 1998; 276:491–504. [PubMed: 9512718]
38. Jamin M, Yeh SR, Rousseau DL, Baldwin RL. J Mol Biol. 1999; 292:731–740. [PubMed: 10497035]
39. Weisbuch S, Gerard F, Pasdeloup M, Cappadoro J, Dupont Y, Jamin M. Biochemistry. 2005; 44:7013–7023. [PubMed: 15865446]
40. Uzawa T, Akiyama S, Kimura T, Takahashi S, Ishimori K, Morishima I, Fujisawa T. Proc Natl Acad Sci U S A. 2004; 101:1171–1176. [PubMed: 14711991]
41. Uzawa T, Nishimura C, Akiyama S, Ishimori K, Takahashi S, Dyson HJ, Wright PE. Proc Natl Acad Sci U S A. 2008; 105:13859–13864. [PubMed: 18779573]
42. Roder H. Proc Natl Acad Sci U S A. 2004; 101:1793–1794. [PubMed: 14769941]
43. Nishimura C, Wright PE, Dyson HJ. J Mol Biol. 2003; 334:293–307. [PubMed: 14607120]
44. Nishimura C, Dyson HJ, Wright PE. J Mol Biol. 2006; 355:139–156. [PubMed: 16300787]
45. Eliezer D, Chung J, Dyson HJ, Wright PE. Biochemistry. 2000; 39:2894–2901. [PubMed: 10715109]
46. Jennings PA, Stone MJ, Wright PE. J Biomol NMR. 1995; 6:271–276. [PubMed: 8520219]
47. Latypov RF, Cheng H, Roder NA, Zhang J, Roder H. J Mol Biol. 2006; 357:1009–1025. [PubMed: 16473367]
48. Pace CN. Methods Enzymol. 1986; 131:266–280. [PubMed: 3773761]
49. Berberan-Santos MN, Martinho JMG. J Chem Educ. 1990; 67:375–379.
50. Fersht, AR. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. 1. Freeman, WH., editor. Cambridge, UK: 1999.
51. Ptitsyn OB, Ting KL. J Mol Biol. 1999; 291:671–682. [PubMed: 10448045]
52. Jamin M, Baldwin RL. Nat Struct Biol. 1996; 3:613–618. [PubMed: 8673605]
53. Samatova EN, Melnik BS, Balobanov VA, Katina NS, Dolgikh DA, Semisotnov GV, Finkelstein AV, Bychkova VE. Biophys J. 2010; 98:1694–1702. [PubMed: 20409491]
54. Walkenhorst WF, Green SM, Roder H. Biochemistry. 1997; 63:5795–5805. [PubMed: 9153420]
55. Sanchez IE, Kiefhaber T. J Mol Biol. 2003; 325:367–376. [PubMed: 12488101]
56. Sauder JM, MacKenzie NE, Roder H. Biochemistry. 1996; 35:16852–16862. [PubMed: 8988024]
57. Hughson FM, Barrick D, Baldwin RL. Biochemistry. 1991; 30:4113–4118. [PubMed: 2021603]
58. Lin L, Pinker RJ, Forde K, Rose GD, Kallenbach NR. Nat Struct Biol. 1994; 1:447–451. [PubMed: 7664063]
59. Chen Y, Barkley MD. Biochemistry. 1998; 37:9976–9982. [PubMed: 9665702]
60. Haruta N, Kitagawa T. Biochemistry. 2002; 41:6595–6604. [PubMed: 12022863]
61. Tcherkasskaya O, Ptitsyn OB. FEBS Lett. 1999; 455:325–331. [PubMed: 10437798]
62. Nishimura C, Dyson HJ, Wright PE. J Mol Biol. 2002; 322:483–489. [PubMed: 12225742]
63. Tanford C. Adv Protein Chem. 1968; 23:121–282. [PubMed: 4882248]
64. Voelz VA, Bowman GR, Beauchamp K, Pande VS. J Am Chem Soc. 2010; 132:1526–1528. [PubMed: 20070076]
65. Nishimura C, Dyson HJ, Wright PE. Proc Natl Acad Sci U S A. 2005; 102:4765–4770. [PubMed: 15769860]

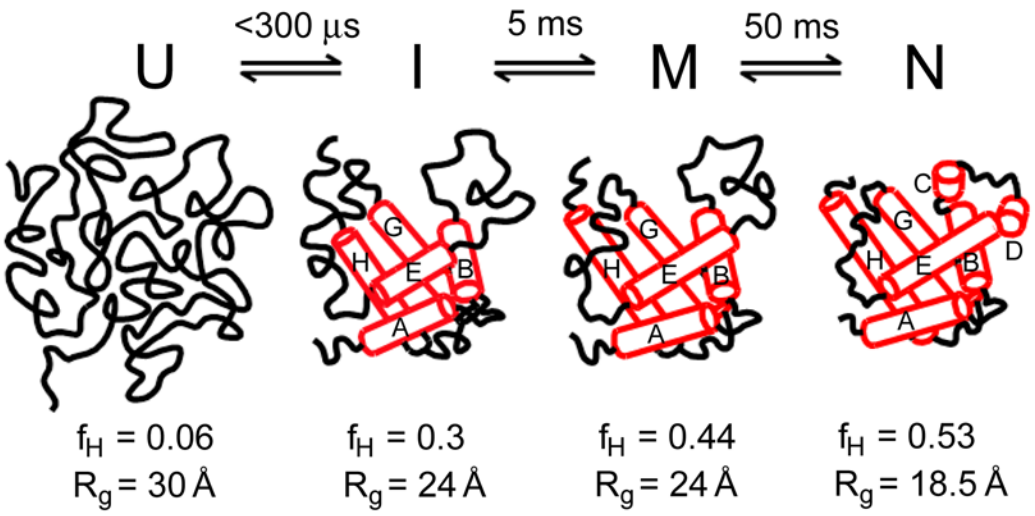

Figure 1.

Folding mechanism of apomyoglobin: a sequential four-state mechanism involving two partially folded intermediate states, I_1 and I_2, in addition to the acid-unfolded (*U*) and native (N) states. The cartoon shows conformational changes accompanying folding, as gauged by the changes in helix content (f_H) and radius of gyration (R_g) reported by Uzawa et al40. Cylinders depict helices and their relative positions in the states along the folding pathway, based on H/D exchange labeling data[65].

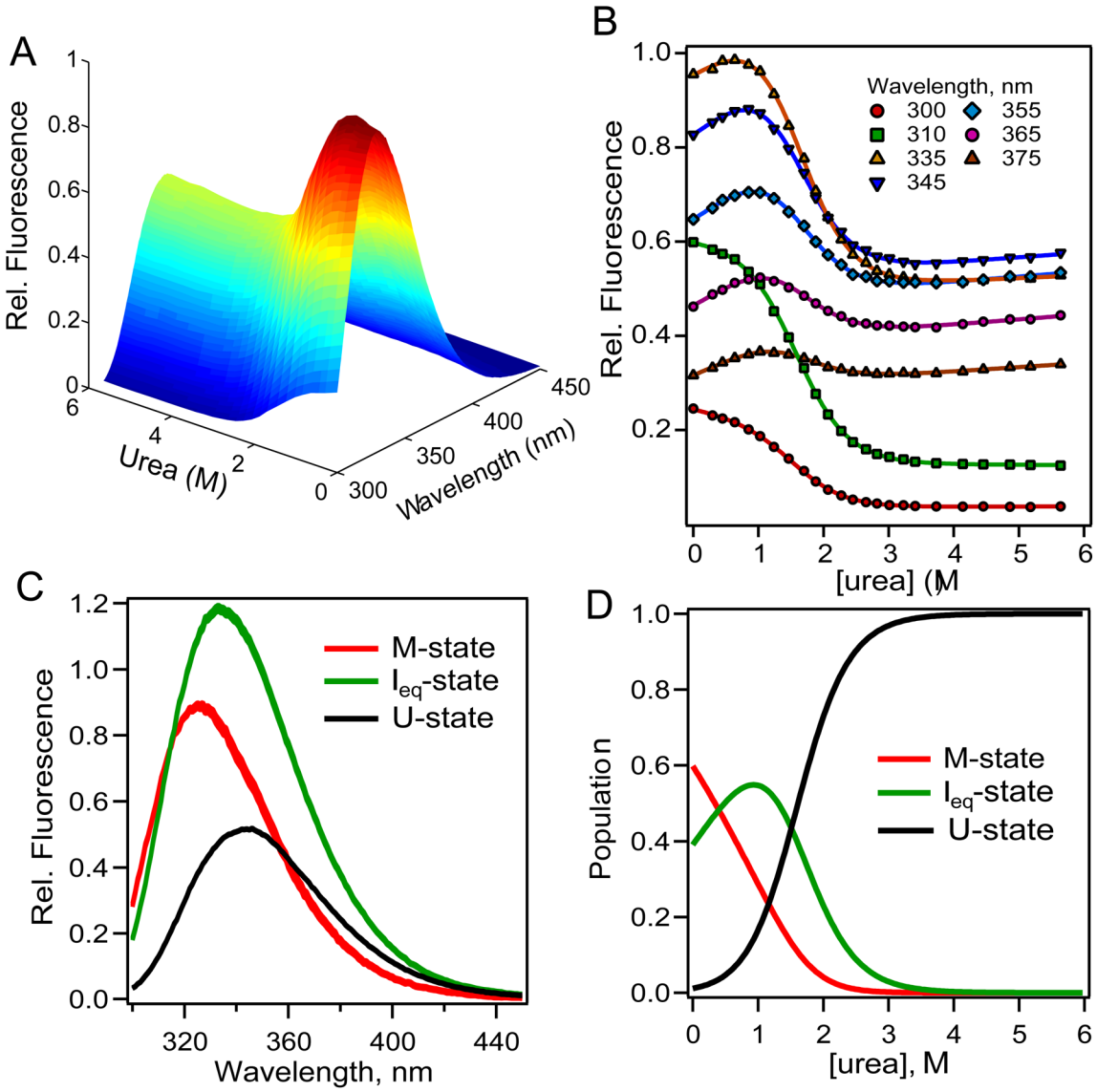

Figure 2.

Global three-state analysis for equilibrium urea-unfolding of apoMb at pH 4.2 monitored by Trp fluorescence spectroscopy. The set of Trp fluorescence emission spectra recorded at different urea concentrations (A) was transposed into unfolding transitions at different wavelengths as a function of urea concentration (B). The transposed dataset was fitted globally to an equilibrium three-state model ($M \leftrightarrow I_{eq} \leftrightarrow U$), as shown by solid lines. Local parameters from the global fitting (y-intercepts unique to individual transitions) yield intrinsic Trp fluorescence spectra for each state at 0 M urea (C). The globally fitted thermodynamic parameters determine the relative populations for each state (D). All titration experiments were conducted at 11 °C.

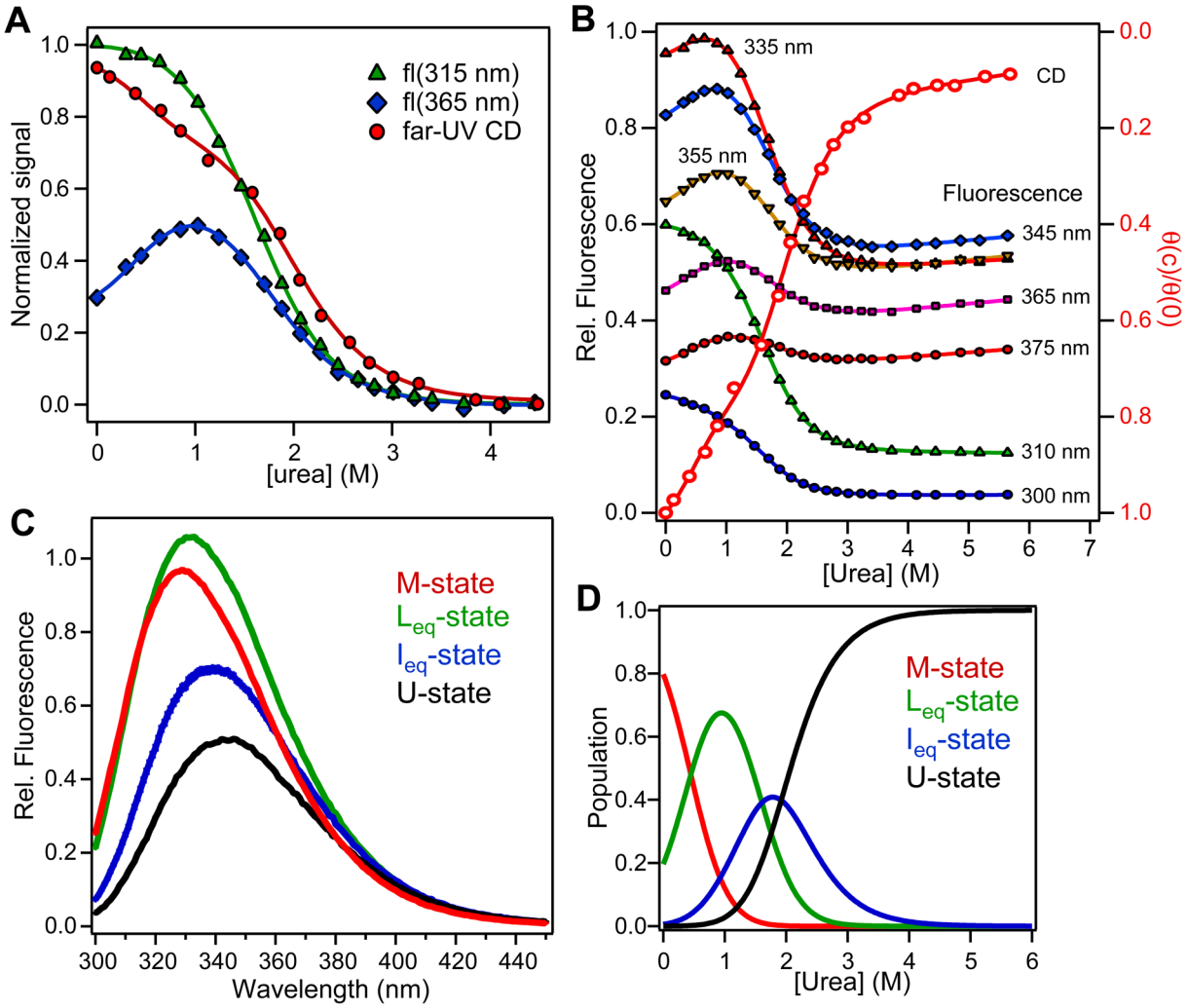

Figure 3.

Global four-state analysis of the combined Trp fluorescence and far-UV CD data of equilibrium urea-unfolding of apoMb at pH 4.2. (A) Normalized urea-dependent fluorescence curves at representative wavelengths (315 nm and 365 nm) are compared with the normalized far-UV CD signal vs. [urea] (red circles). (B) Global four-state fitting of the combined Trp fluorescence (left axis) and far-UV CD data (red circles, right axis). The equilibrium parameters (C_m and m-values for each transition) are listed in Table 1. A constrained fit of the normalized CD data in panel A using these parameters yields relative helix contents of 1, 0.74, 0.62 and 0 for *M*, L_{eq}, I_{eq} and *U*, respectively. (C) Fitted intrinsic Trp fluorescence spectra for each state at 0 M urea. (D) Relative populations of each state calculated from globally fitted thermodynamic parameters.

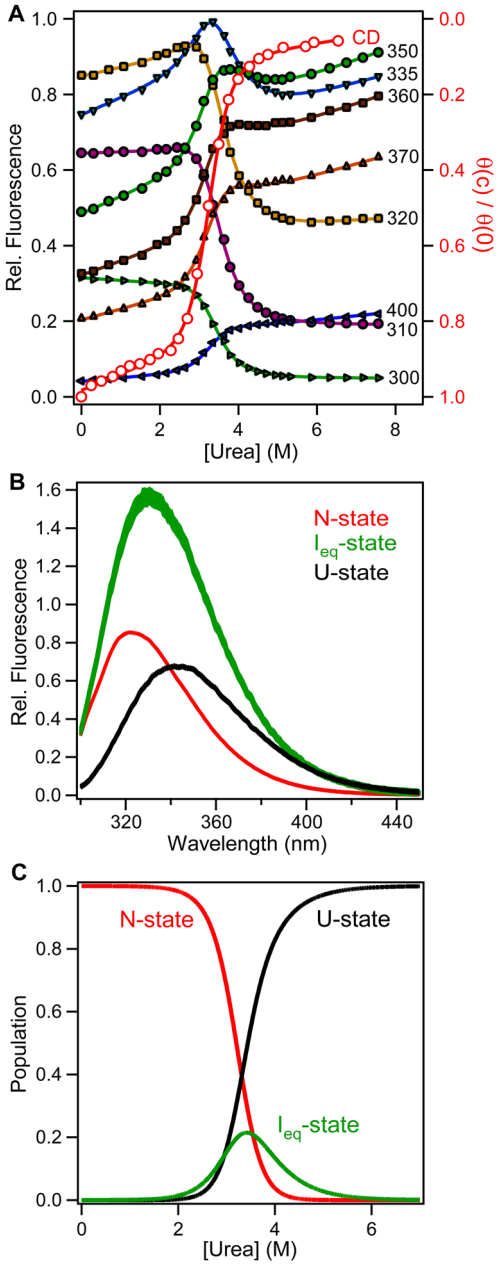

Figure 4.

(A) Global three-state analysis of the combined Trp fluorescence and far-UV CD data of apoMb unfolding at pH 6.2. Urea titration was performed in 20 mM MES buffer, pH 6.2. (B) Fitted intrinsic fluorescence spectra for each state at 0 M urea. (C) Relative populations of each state calculated from globally fitted thermodynamic parameters.

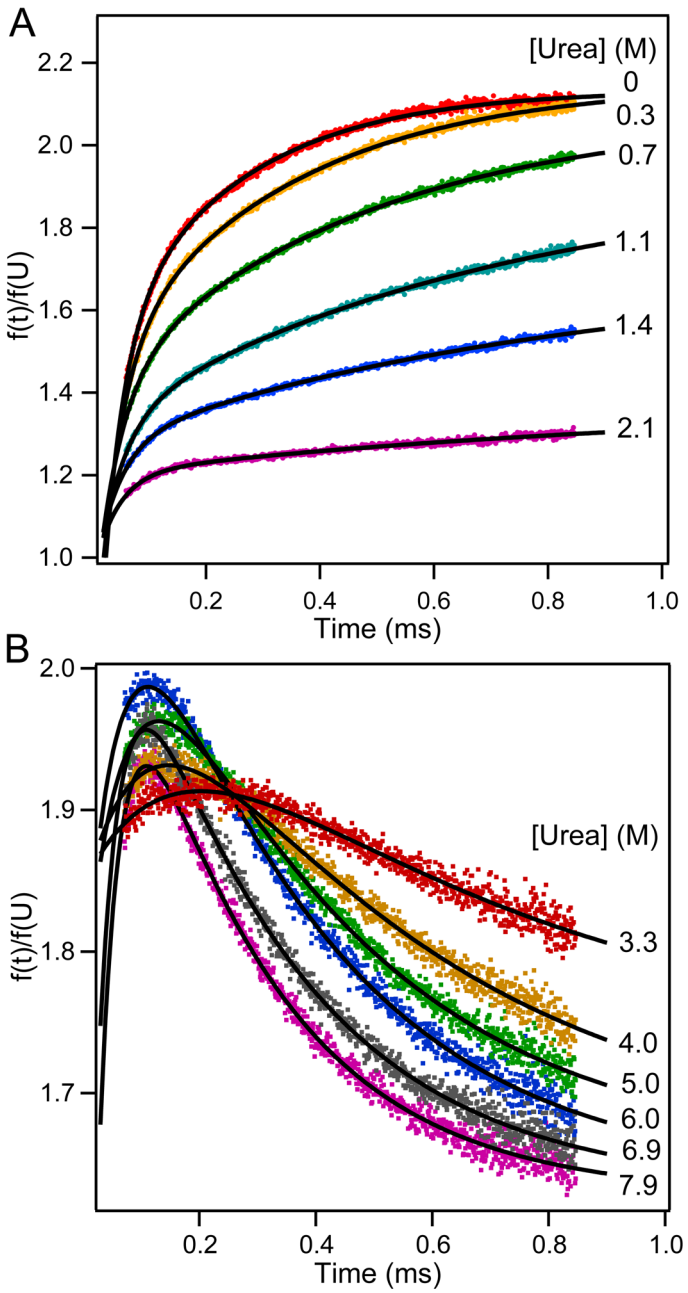

Figure 5.

Kinetics of folding and unfolding of apoMb at pH 4.2. Tryptophan fluorescence changes associated with folding (A) and unfolding (B) of apoMb measured on continuous-flow turbulent mixing apparatus. Solid lines are double exponential fits. For folding (A), the acid unfolded protein (10mM HCl pH 2.0) was mixed with refolding buffer (20 mM sodium acetate with various urea concentration) to achieve the intended urea concentration at a final pH of 4.2. For unfolding (B), protein in 20 mM sodium acetate, pH 4.2 was mixed with unfolding buffer (20 mM NaAc, pH 4.2 with various urea concentrations). The kinetic traces were normalized with the fluorescence signal of the acid unfolded apoMb. All experiments were performed at 11 °C.

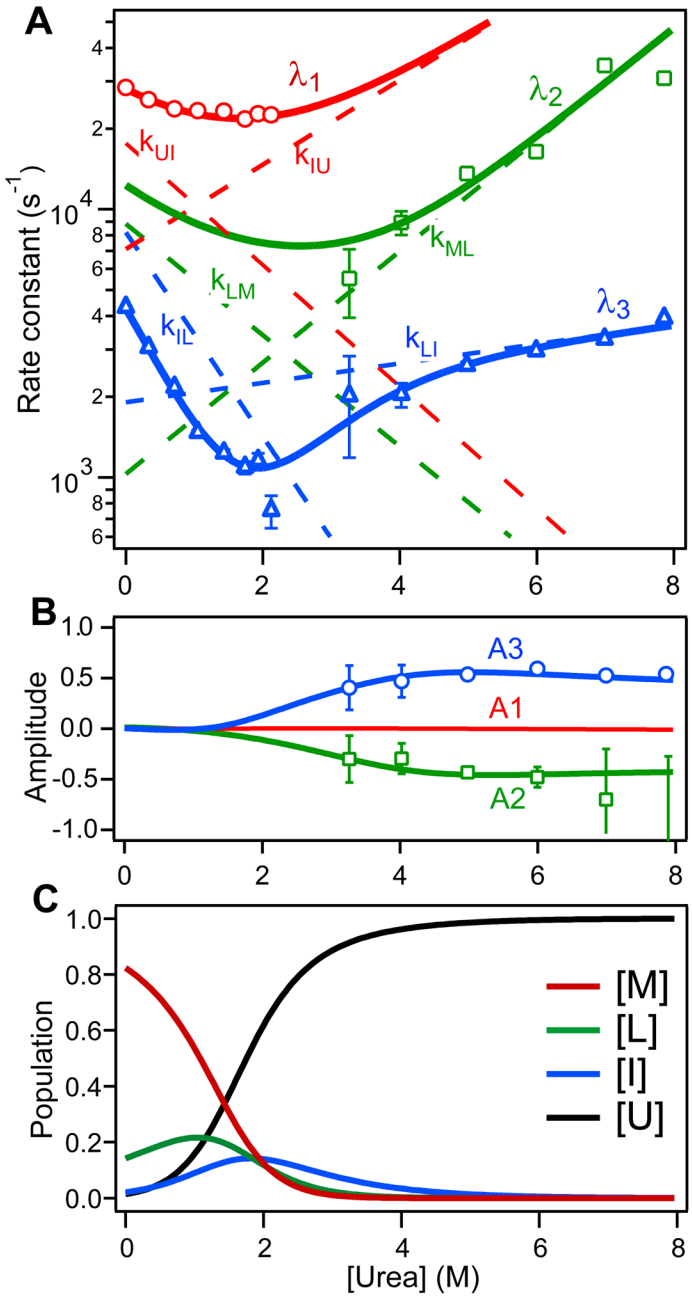

Figure 6.

Four-state kinetic modeling of apoMb folding/unfolding at pH 4.2. (A) Log(rate) vs. [urea] (chevron) plot of observed and predicted rate constants. Circles are apparent rate constants obtained from double exponential fitting of refolding kinetic traces (Figure 5A). Squares are apparent rates obtained from double exponential fitting of unfolding kinetic traces (Figure 5B). Error bars represent fitting errors. Solid lines represent the three observable rate constants, λ_1-λ_3, vs. [urea] predicted by a sequential four-state kinetic model (Scheme A). The corresponding elementary rate constants are shown as dashed lines. The symbols in panels B and C show the relative kinetic amplitudes observed in refolding (B) and unfolding (C) experiments. The lines labeled A1, A2 and A3 represent the predicted amplitudes corresponding to λ_1, λ_2 and λ_3, respectively. Note that the fastest phase, λ1, is predicted to

have zero amplitude during unfolding. The relative amplitude for each state and their urea-dependence are listed in Table 3. (D) Relative population of each state at equilibrium calculated using elementary rate constants and kinetic m-values obtained from the kinetic modeling.

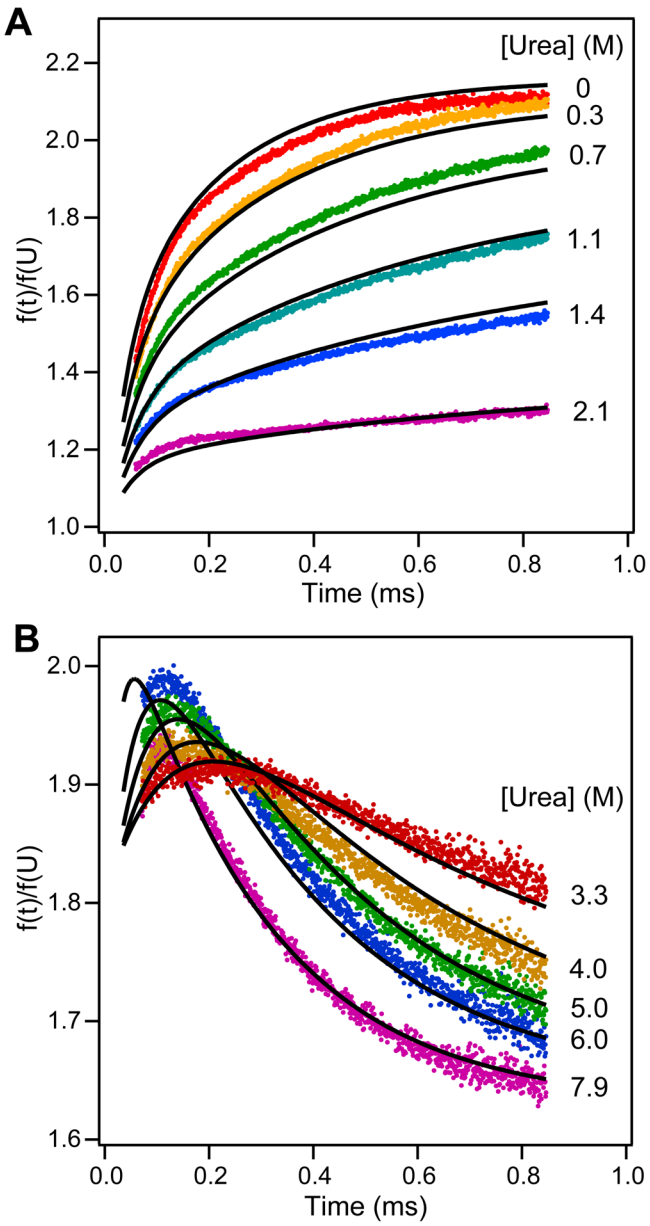

Figure 7.

Global modeling of the family of fluorescence-detected kinetic traces vs. [urea] during folding (A) and unfolding (B) of apoMb at pH 4.2. Kinetic parameters in Table 2 and relative fluorescence yields in Table 3 were used for the direct modeling of the kinetic traces.

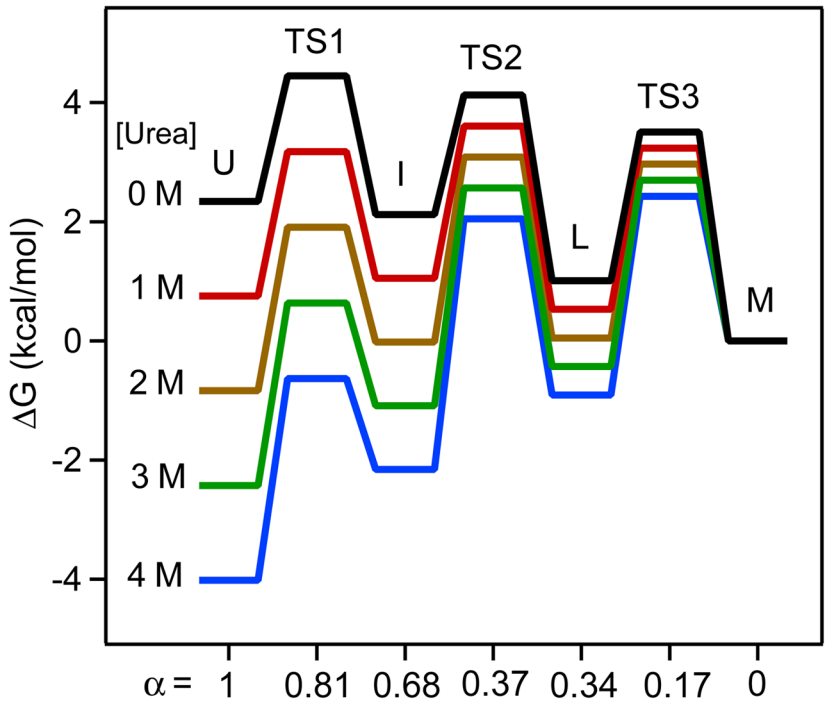

Figure 8.

Free energy diagram for apoMb at pH 4.2. The activation energy for each transition was calculated based on Eq. (3), with an Arrhenius pre-exponential factor of 5 x 10^5 M^{-1} s^{-1}. The α value calculated according to Eq. (4) refers to the relative change in solvent accessible surface area. TS: transition state.

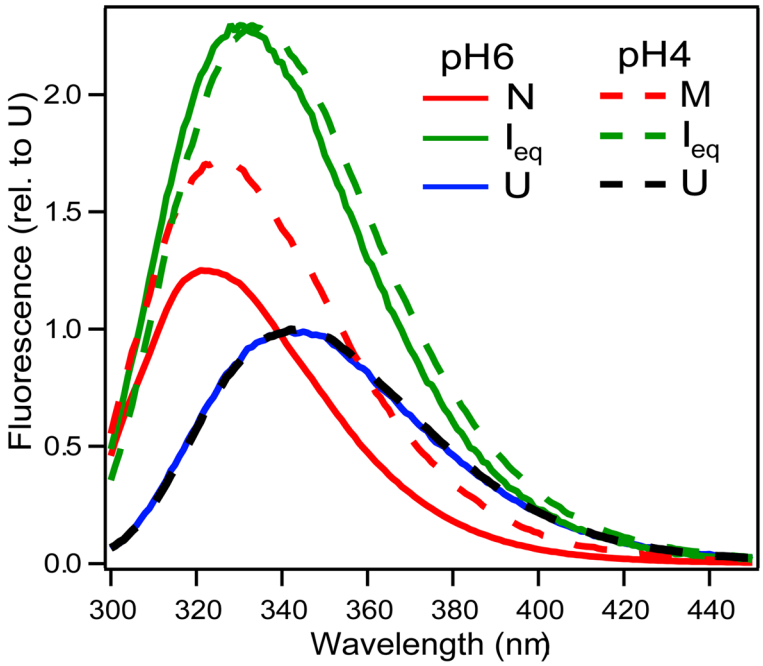

Figure 9.

Fluorescence emission spectra for the predominant equilibrium states of apoMb at pH 4.2 (dashed lines) and pH 6.2 (solid lines), normalized relative to the *U*-state. The spectra were obtained by global three-state fitting (see Figures 2 and 4).

Table 1

Thermodynamic parameters for sperm whale apoMb at pH 4.2 (M-state) and pH 6.2 (N-state) derived by global analyses of urea-induced unfolding monitored by tryptophan fluorescence and far-UV CD spectroscopy using three-state or four-state equilibrium models.

	m (kcal·mol^{-1}·M^{-1})	C_m (M)	ΔG^o (kcal mol^{-1})
3-state model, pH 4.2 [a]			
$M \leftrightarrow I_{eq}$	0.61 ± 0.03	0.40 ± 0.04	0.24
$I_{eq} \leftrightarrow U$	1.32 ± 0.01	1.51 ± 0.01	1.99
4-state model, pH 4.2 [b]			
$M \leftrightarrow L_{eq}$	1.73 ± 0.06	0.43 ± 0.01	0.74
$L_{eq} \leftrightarrow I_{eq}$	1.24 ± 0.04	1.61 ± 0.03	2.00
$I_{eq} \leftrightarrow U$	1.14 ± 0.04	1.93 ± 0.05	2.19
3-state model, pH 6.2 [b]			
$N \leftrightarrow I_{eq}$	1.55 ± 0.01	3.54 ± 0.02	5.50
$I_{eq} \leftrightarrow U$	0.98 ± 0.01	2.95 ± 0.03	2.88

[a] Fluorescence data only.

[b] Combined fluorescence and far-UV CD data

Table 2

Elementary rate constants and kinetic m-values for WT apoMb at pH 4.2 estimated by kinetic modeling of continuous-flow folding/unfolding data, using a four-state mechanism (Scheme B).

Transition (i → j)	$k_{ij}^{\ddagger}$ (s^{-1})	$m_{ij}^{\ddagger}$ ($kcal \cdot mol^{-1} \cdot M^{-1}$)
$U \rightarrow I$	18000	−0.295
$I \rightarrow U$	7000	0.205
$I \rightarrow L$	8500	−0.495
$L \rightarrow I$	1900	0.047
$L \rightarrow M$	9000	−0.27
$M \rightarrow L$	1010	0.27

Table 3

Relative fluorescence yield in the absence of urea (normalized relative to the acid-unfolded state at pH 2) and urea-dependent slope for each state derived by direct fitting of a four-state mechanism to the kinetics of folding and unfolding (Figure 7).

State	Folding		Unfolding	
	Signal	Slope	Signal	Slope
U	0.91	0.05	1.70	−0.01
I	1.85	−0.10	2.12	−0.10
L	2.28	−0.10	2.27	−0.02
M	2.17	−0.16	1.83	−0.03