Fast Events in Protein Folding: Relaxation Dynamics and Structure of the I Form of Apomyoglobin[†]

Rudolf Gilmanshin,[‡] Skip Williams,[§] Robert H. Callender,* William H. Woodruff,[||] and R. Brian Dyer[||]

Department of Physics, The City College of the City University of New York, New York, New York, 10031, Phillips Laboratory/GPID, 29 Randolph Rd., Hanscom AFB, Massachusetts 01731, Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, and CST-4, Mail Stop J586, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received March 19, 1997; Revised Manuscript Received September 23, 1997[⊗]

ABSTRACT: The fast relaxation dynamics of the acid destabilized I form of apomyoglobin (pH* 3, 0.15 M NaCl; apoMb-I) following a laser-induced temperature-jump have been probed using time-resolved infrared spectroscopy. Only a fast, single exponential phase is observed (bleach centered at $\nu = 1633$ cm^{-1} and transient absorbance at 1666 cm^{-1}) with relaxation times of 38 ns at 30 °C and 36 ns at 57 °C; no additional slow (microsecond) phase is observed as previously found in the native form of apomyoglobin. Folding times of approximately 66 ns are derived from the observed rates based on a simple two-state model. The equilibrium melting of the 1633 cm^{-1} component shows noncooperative linear behavior over the temperature range studied (10−60 °C). The low amide I′ frequency, the fast relaxation dynamics, and the noncooperative melting behavior are characteristic of isolated solvated helix. The analysis of the amide-I′ band reveals another major component at 1650 cm^{-1} assigned to native-like structure stabilized by tertiary contacts involving the AGH core, which does not show dynamic or static melting under our conditions. ApoMb-I has generally been taken to be a "molten globule" species. The present results indicate a heterogeneous structure consisting of separate regions of native-like unit(s), solvated helices, and disordered coil, excluding a homogeneous molten globule as a model for apoMb-I. From the current studies and other results, a detailed model of the folding of apomyoglobin is presented.

It is becoming increasingly apparent that the earliest events in protein folding are critical to understanding the mechanism of the folding process. Unfortunately, the transition from the distribution of unfolded states to more compact structures has not yet been resolved, because such a transition may be intrinsically fast and because of the relatively slow nature of conventional reaction initiation techniques (such as stopped-flow) and the temporal limitations of experimental probes which are sensitive to polypeptide structure (Evans & Radford, 1994; Kim & Baldwin, 1990; Schmid, 1992). Obviously, structural detail on the mechanisms of this transition can only be ascertained by fast, structure-specific measurements. Our primary probe for polypeptide structure is the amide I infrared absorbance, which consists to a good approximation of the superposition of the amide C═O stretching vibrations. Amide I has well-established sensitivity to secondary structure and also gives information about tertiary structure (Haris & Chapman, 1995; Gilmanshin et al., 1997b; Susi & Byler, 1986). In our experiments, the folding/unfolding relaxation kinetics are initiated by a laser-induced temperature-jump. The short-time limit of the measurements reported here, set by laser pulse duration and detector risetime, is ca. 10 ns.

Myoglobin contains eight strands of mostly α-helical segments, labeled A−H (Evans & Brayer, 1988). Apomyoglobin is prepared by removing the heme group from myoglobin. The most stable form of the apomyoglobin, formed near neutral pH, adopts a structure that is native-like according to the available NMR, CD, calorimetric, and hydrodynamic evidence (Griko et al., 1988; Hughson et al., 1990; Johnson & Walsh, 1994; Lecomte et al., 1996). This form, called herein apoMb-N,[1] has a tightly packed core, consisting of at least the very stable A, G, and H helices, with roughly the same secondary structure and tertiary fold as the holoprotein. The major differences appear to occur in the C, D, and F helices, which are considerably more disordered in the apoprotein.

Previously, we reported the time scales associated with the dynamics of secondary and tertiary structure formation in the native like form of apomyoglobin. In that study (Gilmanshin et al., 1997b), we found that, despite the relative structural simplicity of this protein, apoMb-N consists of substructures which have different thermodynamic stabilities and kinetic labilities. In particular, we found that apoMb-N contains a region of α-helical structure that melts with a cooperative transition, in agreement with previous results, and which shows folding kinetics on the 200−300 μs time scale at the ca. 60 °C temperature reached in the temperature-jump (T-jump). We also observed a region of α-helices which are exposed to solvent and which form their secondary

[†] This work was supported by the National Institutes of Health, GM53640 (R.B.D.) and GM35183 (R.H.C.), and the National Science Foundation, MCB-9417892 (R.H.C.).

* Corresponding author at the Einstein College of Medicine, on leave from City College of New York. Fax: 212 650 5503. E-mail: call@sci.ccny.cuny.edu.

[‡] City College of New York.

[§] Philips Laboratory.

[||] Los Alamos National Laboratory.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1997.

[1] Abbreviations: apoMb, horse apomyoglobin; apoMb-N, native form of apoMb; apoMb-I, an acid form of apoMb which occurs at pH <4 and which requires excess of neutral salt to stabilize its structure at lower pH; T_m, midpoint temperature of the melting transition; λ_{max}, maximum position of fluorescence spectrum.

S0006-2960(97)00634-X CCC: $14.00

structure with a lifetime of approximately 100 ns. The latter lifetime is similar to that which we measured earlier for a model α-helical 21-peptide (Williams et al., 1996), wherein the formation of solvated α-helix occurs with a lifetime of about 180 ns (27 °C). In addition, apoMb-N contains helices which did not melt under the conditions that we studied (Gilmanshin et al., 1997a,b). Therefore, the native form of apoMb consists of three regions of folded substructures of sharply contrasting dynamical natures: two helical substructures which contain tertiary contacts and exhibit cooperative melting behavior, but with different stabilities (i.e., different melting T_ms) and kinetic labilities, and a portion of solvated helices that make little or no stabilizing contacts.

Here, we report time-resolved infrared measurements of the folding dynamics for the "I" form of apomyoglobin (apoMb-I) as well as characterization of its structure by FT-IR and fluorescence spectroscopies. The I form is reached by lowering pH, specifically to pH 4 (Griko et al., 1988). It is also stabilized at lower values of pH by the addition of salt (Fink et al., 1991). Many proteins exhibit a "burst phase" in stopped-flow folding experiments, in which the development of spectroscopic (e.g., CD and fluorescence) signatures of a collapsed state with secondary structure occurs within the dead time the experiments (1−10 ms) (Evans & Radford, 1994; Kim & Baldwin, 1990). Often, the burst phase state can also be substantially populated under certain equilibrium conditions (Ptitsyn, 1992). The I state of apoMb has been observed to share many structural features with a species that forms within the burst phase of stopped-flow measurements as ascertained by CD, hydrogen/deuteron exchange, and small angle X-ray scattering measurements (Jennings & Wright, 1993; Eliezer et al., 1995). It is therefore a model of an early folding intermediate. Many studies have suggested apoMb-I is "molten globule"-like (but see below). That is, it retains considerable secondary structure and is quite compact, similar to the native state of the protein. However, molten globules are generally assumed to have fluctuating structures and show noncooperative melting because they lack substantial fixed tertiary interactions found in compact native structures. It is clear that the relaxation dynamics of apoMb-I are of great interest in elucidating the mechanisms of folding as is the exact nature of its structure.

MATERIALS AND METHODS

Materials. ApoMb was prepared from horse heart myoglobin (Sigma) by a modified 2-butanone extraction of the heme group (Rothgeb & Gurd, 1978; Teale, 1959). The extraction was followed by exhaustive dialysis of the protein solution against 0.01 M HCl and then against 0.01 M sodium phosphate buffer with 0.04 M NaCl, pH 6.0. The sample was clarified by centrifugation and filtered on a Sephadex G75 column (Pharmacia Biotech). Monomeric apoMb fractions were extensively dialyzed against deionized water and lyophilized. All operations were performed at 4 °C. Final sample contained less than 0.3% of holomyoglobin and was completely homogeneous on SDS−PAAG electrophoresis. The absorption at 280 nm ($E_{1\%} = 8.4$ cm^{-1}) was used to determine concentration (Crumpton & Polson, 1965).

All solutions were prepared with D_2O and contained 0.15 M NaCl at pH* 3.0. The pH* is uncorrected pH-meter reading at 20 °C. The pH* was controlled after the protein dissolution and adjusted if needed. The apoMb solution was equilibrated against reference solvent using a small Sephadex G25 column for better compensation of residual H_2O in the IR measurements. Sample concentrations of 0.08 mg/mL were used for fluorescence measurements and 4−7 mg/mL for the IR measurements. The protein tended to aggregate at concentrations of 4 mg/mL and larger during equilibrium measurements at temperatures higher than 55 °C. However, no aggregation was observed during kinetics measurements, even with T-jumps to temperatures higher than 55 °C, suggesting that the aggregation time scale is significantly longer than the millisecond lifetime of the T-jump.

Infrared Spectroscopy and Temperature-Jump Generation. The static infrared spectra were obtained using a Bio-Rad model FTS 60A FTIR spectrometer as described previously (Williams et al., 1996). Teflon spacers of 50 and 100 μm were used for equilibrium and kinetics measurements, respectively. The fundamental aspects of the transient IR spectrometer and the temperature-jump generation have been previously described (Williams et al., 1996; Causgrove & Dyer, 1993). In brief, the apparatus is capable of resolving real-time, single wavelength infrared absorption transients from 20 ns to arbitrarily long times (typically 10 ms) in one laser shot. A widely tunable CW lead salt infrared diode laser functions as the probe which yields IR from 1700 to 1632 cm^{-1}. An injection-seeded, Q-switched Nd:YAG laser, pulsed dye laser, and WEX frequency differencing module produce the pump radiation at 2 μm (10 ns FWHM Gaussian pulse width); this light is absorbed by a weak D_2O IR band (instantaneously on the nanosecond time scale) and is the source for the temperature-jump. The combined total instrument response time is approximately 23 ns, which is currently limited by the fall time of the detector/preamplifier system. Transient changes in the transmission of the CW IR beam through the sample are detected by a HgCdTe detector, digitized, and signal averaged. A split IR cell was employed, one compartment containing D_2O plus protein and the other containing D_2O. The cell was translated so that each compartment was probed under identical conditions.

Fluorescence Measurements. Fluorescence spectra were measured on FluoroMax-2 spectrofluorimeter (Instruments S. A., Inc.) with correction for a spectral dependence of registration response. The excitation wavelength was 290 nm, with a 2.5 nm band-pass for both excitation and emission. A $1 \times 1 \times 4.5$ cm^3 cuvette was used. The data were collected with DM-3000 software (Instruments S. A.) and processed with IgorPro (WaveMetrics, Inc.) software. Solvent background was measured at the same conditions and subtracted from every protein fluorescence spectrum.

RESULTS

Equilibrium-Melting Behavior. The static, equilibrium unfolding of apoMb was followed by obtaining the FTIR spectra in the amide I′ region (Figure 1; the prime denotes the frequencies of deuterated amide groups). The amide I′ absorption envelope is centered at 1645 cm^{-1} at low temperature but broadens and shifts to higher frequency as temperature increases. Resolution enhancement of the equilibrium FTIR spectra via second derivative analysis (Figure 2) reveals two major amide I′ subcomponents at 1650 and 1633 cm^{-1}. The considerable overlap of the subcomponents in the absorbance spectra (Figure 1a) hinders identification of the equilibrium changes produced by thermal denaturation. A simple difference analysis highlights these changes (Figure 1b); as the temperature is increased, a trough

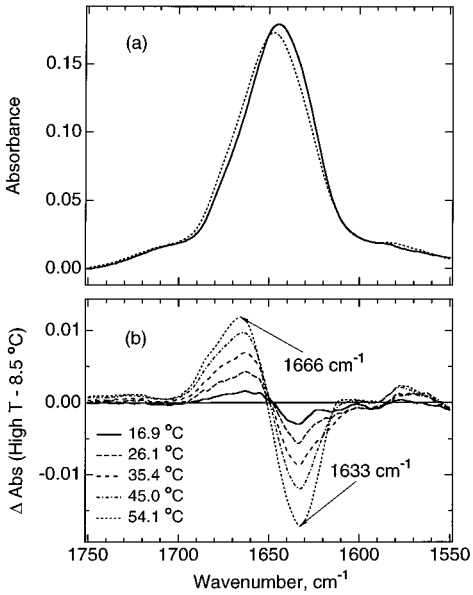

FIGURE 1: Equilibrium FTIR absorbance spectra of apoMb in D_2O (4.8 mg/mL in 0.15 M NaCl solution, pH* 3.0) in the amide I′ spectral region, as a function of temperature. The temperature dependent D_2O background has been subtracted from each FTIR spectrum. (a) Comparison of the spectra at 8.5 (—) and 54.1 °C (···). (b) Difference spectra, generated by subtracting the spectrum at 8.5 °C from the spectrum at each temperature.

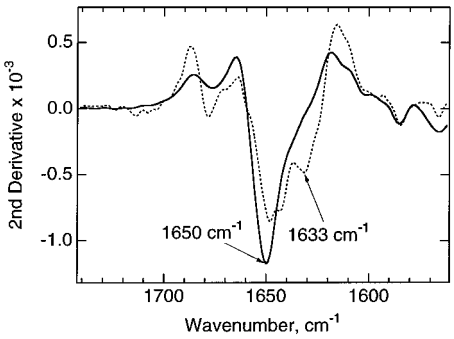

FIGURE 2: Comparison of the second derivative spectra for apoMb-N (solid line) and apoMb-I (dashed line) at 8.5 °C. The second derivatives were smoothed with a fourth order, 19 point Savitzky-Golay algorithm (corresponds approximately to 20 cm^{-1} band-pass for our dataset).

appears in the amide I′ difference at 1633 cm^{-1} with a concomitant growth of a peak at 1666 cm^{-1}. These observations serve as a guide for band assignments and for the kinetics experiments.

The amide I band is a well-established indicator of secondary and tertiary structure (Haris & Chapman, 1995; Gilmanshin et al., 1997b; Susi & Byler, 1986); hence, these spectral changes can be assigned to specific structural changes that accompany thermal denaturation. The structural sensitivity of amide I arises from the dependence of the C═O stretching frequency of the polypeptide backbone carbonyls on hydrogen bonding and on structure-dependent vibrational coupling. Previous studies have shown that in D_2O, disordered coil structures generally exhibit amide I′ frequencies between 1665 and 1675 cm^{-1} (Chirgadze et al., 1973; Reinstädler et al., 1996; Williams et al., 1996), while the characteristic frequency range is 1650−1655 cm^{-1} for "native" α-helix protected from solvent by the tertiary fold (Susi & Byler, 1986; Jackson & Mantsh, 1995). More recent studies (Haris & Chapman, 1995; Reisdorf & Krimm, 1996; Williams et al., 1996) have shown that, in the absence of

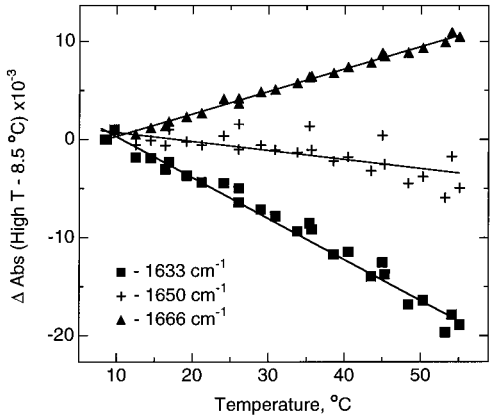

FIGURE 3: The absorbance difference Abs(T) − Abs(8.5 °C) versus temperature of the major subcomponents of the amide I band of apoMb-I, at 1650 cm^{-1} (+, native helix), 1633 cm^{-1} (■, solvated helix), and 1666 cm^{-1} (▲, disordered coil).

protection by tertiary structure, a helix solvated in D_2O exhibits amide I′ frequencies that are considerably lower, near 1632−1645 cm^{-1} (the "solvated" helix).

From these studies and by analogy with our previous results on apoMb-N (Gilmanshin et al., 1997b), the 1633 cm^{-1} component is assigned to solvated helix, the 1650 cm^{-1} component to α-helix protected from solvation by tertiary interactions and the 1666 cm^{-1} peak observed in the difference curves of Figure 1b to random coil. A comparison of the second derivative spectra of apoMb-N and apoMb-I is shown in Figure 2. It is clear from these spectra that the solvated helix component at 1633 cm^{-1} is significantly larger in the apoMb-I form, at the expense of the 1650 cm^{-1} native helix component. This observation is consistent with X-ray and light scattering results which indicate that the I form of apoMb swells in volume by a factor of 1.7−2.4 (an increase in volume equivalent to 570−800 water molecules) compared to the native form (Gast et al., 1994).

The temperature dependences of the major components of the amide I′ band are plotted in Figure 3. The changes in amide I′ intensities are obtained from the difference spectra and represent the effects of temperature on the relative populations of the solvated helical, native helical, and random coil structures to which the spectral components are assigned. The thermal denaturation of the solvated helical component (loss of signal at 1633 cm^{-1} and gain of signal at 1666 cm^{-1}) does not show a sharp transition with temperature, instead exhibiting a continuous loss of structure over the entire temperature range studied. This noncooperative melting is similar to that observed previously for short peptides and for the solvated helix component of apoMb-N (Gilmanshin et al., 1997b; Williams et al., 1996). In contrast, the subcomponent at 1650 cm^{-1} assigned to native helix shows essentially no change in intensity below 55 °C. Thus, by the amide I′ criterion, the solvent-protected helices do not melt to any significant degree over the temperature range of our experiments (10−55 °C). The structural transitions of apoMb-I over the entire temperature regime studied are dominated by the melting of solvated helix.

The FTIR observations do not allow us to assign which part of the protein structure contains the thermally stable substructure that gives rise to the 1650 cm^{-1} infrared band. However, measurements of intrinsic fluorescence spectra allow us to identify this (at least, partially) with the "core" composed of the A, G, and H helices [see also Gilmanshin et al. (1997a)]. Myoglobin contains two tryptophan residues

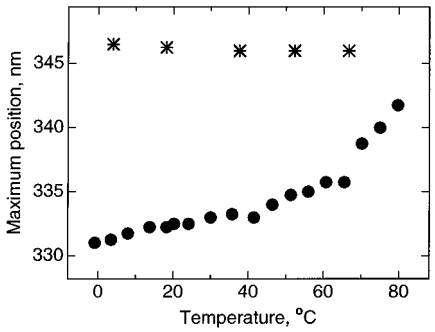

FIGURE 4: Maximum position λ_{max} of the intrinsic fluorescence spectrum of apoMb-I (●) and that of L-tryptophan in D_2O (*) at different temperatures.

and both of them are in the A helix. For these residues to be protected from solvation, the hydrophobic tertiary structure involving the AGH core must be largely intact. Figure 4 shows that this is the case. While many effects may quench the intensity of tryptophan fluorescence, the wavelength of the fluorescence maximum is quite specifically sensitive to polarity of environment. A hydrophobic environment leads to fluorescence λ_{max} near 332 nm, while an aqueous environment engenders a pronounced red-shift with the limit λ_{max} near 347 nm (see λ_{max} for L-tryptophan solution in D_2O, Figure 4). Figure 4 shows that λ_{max} for apoMb-I remains near 332 nm until the temperature is above 50 °C, clearly demonstrating hydrophobic protection for the A helix and thereby an essentially intact AGH core.

Kinetics. The approach to measuring the fast events in the folding of apoMb-I was similar to that reported previously for apoMb-N and model peptides (Gilmanshin et al., 1997b; Williams et al., 1996). We utilize a pulsed laser tuned to a weak absorption band of D_2O to create a laser-induced temperature-jump and initiate the unfolding of apoMb. The time scale of the T-jump is set by the laser pulse width, ca. 10 ns, and temperature-jumps of the magnitude 10−40 °C can be attained within the 10 ns duration of the pulse. The kinetics subsequent to the temperature-jump was followed by transient absorption of infrared diode laser light tuned to frequencies within the amide I′ envelope. The observed relaxation is for the establishment of the new equilibrium, corresponding to a more highly unfolded equilibrium following the temperature-jump, but the relaxation kinetics (given the model discussed below) contain contributions from both folding and unfolding. T-jumps from different starting temperatures (e.g., jumps from 10 to 30 °C and from 45 to 57 °C) were employed to probe within the temperature range investigated in the equilibrium FTIR and fluorescence measurements described above.

Representative kinetics traces in the amide I′ region are shown in Figure 5 for vibrational frequencies within the absorption envelopes which correspond to changes in the solvated helix (absorbance bleach at 1633 cm^{-1}) and the disordered structure (positive absorbance transient at 1670 cm^{-1}). Only one kinetics relaxation is observed between $t = 0$ and 1 ms, and that relaxation is on the time scale of tens of nanoseconds. Single exponential fits to the data at 30 °C convolved with the instrument response function yield a time constant of 40 ns for the bleach and 36 ns for the positive transient absorbance. These two time constants can be considered to be the same within our signal to noise, and an average time of 38 ns is used in what follows below. Similarly, a relaxation time of 36 ns is found for the kinetics

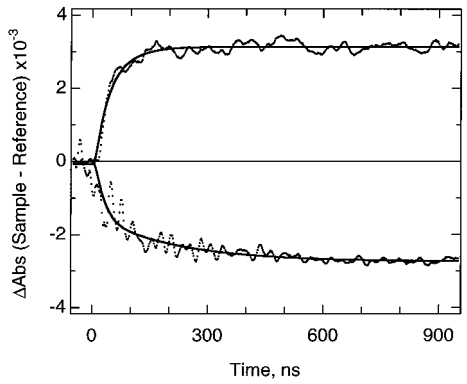

FIGURE 5: IR kinetics response for a temperature-jump from 10 to 30 °C at 1670 cm^{-1} (transient absorbance; near the peak of the positive absorbance in the static difference FTIR spectrum) and 1632 cm^{-1} (transient bleach; near the peak of the bleach in the difference spectrum). The background D_2O response has been subtracted from these data. The solid lines are single exponential fits convoluted with the instrument response.

of the temperature-jump from 45 to 57 °C. In both temperature-jump experiments, both the bleach and positive transient signals begin to show recovery to their baseline values on the millisecond time scale (data not shown), due to heat flow out of the laser interaction volume, with a ca. 2 ms time constant, which returns the sample to its initial temperature (10 or 45 °C, respectively). Similar traces obtained at the native helix frequency (1650 cm^{-1}, not shown) show no change on any time scale. An important distinguishing feature of the kinetics of apoMb-I, therefore, is the loss of the slow (130 μs) kinetics phase centered at 1650 cm^{-1} that was observed for apoMb-N. The disappearance of the slow kinetics phase, together with the loss of native helix intensity observed in the static data, indicate that the native structure that gives rise to this slow phase in apoMb-N is no longer present in the I form.

The correspondence between the static FTIR difference spectra for a given temperature change and the transient absorbance spectra for the same temperature-jump in the kinetics experiments is shown in Figure 6, which compares the transient differences after the relaxation (measured 300 ns after the T-jump) to the FTIR difference spectra. As is the case with the static difference spectra, the transient absorbance spectra are dominated by a bleach at 1633 cm^{-1}, indicating loss of solvated helix, and a positive absorbance at 1670 cm^{-1}, indicating appearance of random coil. The 300 ns spectra at all temperatures reproduce the static difference spectra, and no further evolution of the spectra is observed.

The random coil structure indicated by the infrared absorbance at 1666 cm^{-1} is undoubtedly composed of a heterogeneous mixture of unfolded structures. The gradual melting behavior of the solvated helix of apoMb-I with increasing temperature along with the observed fast kinetics of interconversion of random and solvated helical structures [similar to that of the "Fs" 21-peptide that we have previously characterized (Williams et al., 1996)] suggests a low energy transition barrier separating these structures. Furthermore, the rate is essentially temperature independent over the range studied, also supporting a small transition barrier. Thus, the solvated helical portion of apoMb-I (even with its helices intact) almost certainly exists as a heterogeneous population. The observed single exponential kinetics indicate that either the heterogeneous populations of both folded and unfolded

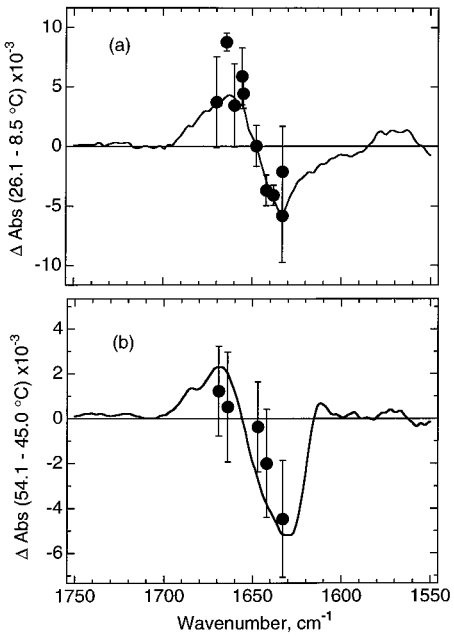

FIGURE 6: (a) FTIR difference spectrum (26.1−8.5 °C) for apoMb-I (solid line) compared with the time-resolved infrared spectrum at 300 ns (circles) after the T-jump; (b) FTIR difference spectrum (54.1−45.0 °C) for apoMb-I (solid line) compared with the time-resolved infrared spectrum at 300 ns (circles) after the T-jump.

states are in rapid equilibrium, or that each state has a similar barrier to melting. There are convincing theoretical and experimental reasons to believe that these states are in rapid equilibrium (Schwartz, 1965; Gruenewald, et al., 1979; Zwanzig, 1995; Eaton et al., 1996). The rate of interconversion between heterogeneous helix structures and random sequence can be cast in a Zimm−Bragg approach and involves a number of parameters, for example, the propensity of helix nucleation and growth and the formation rate of an additional single loop of helix [see, e.g., Schwartz (1965)]. Moreover, any theory to understand the solvated helix-random structure interconversion in apoMb-I must take into account the possibility of helix−helix interactions. Therefore, a rigorous analysis of the kinetics requires a structural model of the solvated helical portion of apoMb-I which is not now available. The calculations of Schwartz (1965), simplified by taking the limit structure to be a homogeneous long polypeptide, yielded an approximate relaxation time of 100 ns for the helix−coil transition, which is perhaps surprisingly close to that determined for apoMb-I. We have now observed essentially the same relaxation time constants for solvated α-helix/disordered structure interconversion in three other systems: the synthetic Fs helical peptide Suc-$A_5(A_3RA)_3A$-NH_2 [relaxation time of 160 ns at 27 °C (Williams et al., 1996)]; the S-peptide of ribonuclease A [relaxation time of <300 ns at 25 °C (Callender et al., 1994)]; and the solvated portion of apoMb-N [48 ns at 60 °C (Gilmanshin et al., 1997b)]. The similarity of the relaxation times for these very different systems suggests that it is a typical, intrinsic relaxation time for the solvated helix of intermediate length.

Given that the T-jump relaxations show single exponential kinetics and that there is rapid equilibrium between the two states, it is useful to approximate the results as a two-state process between the distributions of folded and unfolded structures. In this simplified case, the observed relaxation obeys a reversible first-order rate law wherein $1/\tau_{obs} = 1/\tau_f + 1/\tau_u$, ($\tau_{obs}$ is the observed relaxation time, τ_f and τ_u are the folding and unfolding life times, respectively, where both τ_f and τ_u are averaged for the ensembles of substructures) and the equilibrium constant favoring folding $K_f = \tau_u/\tau_f$. Using the relative intensities at the marker positions of the solvated (1633 cm^{-1}) and disordered (ca. 1666 cm^{-1}) components of amide I′ as proportional to their composition, K_f is 1.4 at 30 °C and 1.2 at 57 °C. These values for K_f must be considered as very approximate, although probably better than 1 order of magnitude (the calculated value is likely a lower limit; however, the temperature dependence at these frequencies is a maximum). Using these values and the measured relaxation time of τ_{obs} of 38 ns at 30 °C and 36 ns at 57 °C yield folding times, τ_f, of 65 ns and 66 ns at the respective temperatures.

DISCUSSION

In relaxation experiments such as here, it is very important that the structure of the unfolded as well as the folded states be well characterized. It is difficult to obtain a completely unfolded protein apart from the highest extremes of denaturing conditions, and changes in temperature generally do not fully unfold a protein. Moreover, the kinetics of folding and, particularly, unfolding are typically dominated by one rate-limiting step so that other key steps can go unobserved. Under these conditions, a fruitful approach in understanding folding dynamics is to characterize the structures and relaxation kinetics of successively more denatured protein states. In this way, the folding pathway can be "peeled away", and processes on both fast and slower time scales determined. Previously, we reported on the relaxation dynamics of the native state of apomyoglobin, apoMb-N, as induced by fast T-jump studies in experiments very similar to those herein for apoMb-I, the acid denatured form of the protein. Here, we study the I form of apomyoglobin, which is a model for an early intermediate along its folding pathway (see the introductory portion of the paper). In the next section, the folding dynamics as induced by our T-jump experiments as well as the structures of the folded and temperature denatured states as characterized by our IR and fluorescence measurements and studies by others are discussed. In the last section, the folding pathway of apomyoglobin is discussed as deduced by combining the kinetic studies of apoMb-N with those of apoMb-I. It is now possible to give a quite detailed picture of the folding of apomyoglobin.

Folding Dynamics of apoMb-I. The acid form of apoMb exhibits simple relaxation behavior when perturbed by a T-jump, yielding a fast, single exponential kinetics phase in the amide I′ IR spectrum. This result by itself seems to support the picture of a homogeneous structure having no fixed tertiary interactions and, therefore, no slow kinetics phase such as that observed for apoMb-N. Superficially, the result fits very well with the view of the acid form of apomyoglobin as a molten globule. A closer examination of the present infrared and fluorescence evidence as well as other studies, however, reveals a much different picture of the structure and dynamics of this form of apoMb.

In the first place, it is clear that apoMb-I contains a native-like substructure. Recently, Kay and Baldwin (1996) showed that the A, G, and H helices, or parts of them, form a native-like subdomain, the AGH core, within the acid form of apoMb by studying urea denaturation of apoMb-I mutants. Moreover, temperature-induced denaturation transitions of different structural forms of apomyoglobin were studied

monitoring intrinsic tryptophan fluorescence (Gilmanshin et al., 1997a). It was found that the tryptophans, which are only located in the A helix, are effectively screened from solvent in the acid form throughout most of the temperature range tested (see also Figure 4). At high temperatures and under stronger destabilizing conditions, however, the tryptophans' fluorescence parameters show sigmoidal thermal denaturation demonstrating that native-like contacts exist in apoMb-I. This conclusion is in agreement with an analysis of the melting behavior exhibited in the present equilibrium FTIR measurements. A native-like substructure is clearly identified and consists of an α-helical region, wherein the peptide carbonyl groups are protected from solvation by native or native-like tertiary contacts. This result arises from the observation of an amide I′ marker component at ca. 1650 cm^{-1}, which is expected for buried helices on the basis of empirical studies of native proteins. Thus, the structure of apoMb-I is anchored at each end of its primary sequence by a compact region, the AGH core, whose contacts are native-like by both the IR and fluorescence criteria as well as the studies of Kay and Baldwin (1996).

The remainder of the apoMb-I sequence consists of a mixture of solvated α-helical and random structure. They are distinguished from one another and from the native-like helical structure by their amide-I′ frequencies (ca. 1633 cm^{-1} for the solvated helix and >1660 cm^{-1} for random). Only the solvated helix content is observed to melt appreciably, whereas the native helix remains essentially unperturbed over the temperature range that we have performed the T-jump measurements (up to ca. 55 °C; Figure 3). The temperature dependence of the IR absorbance for the solvated helix conformation is nearly linear over a wide temperature range which suggests a low cooperativity of melting and a small size of the cooperative unit [see, e.g., Uversky and Ptitsyn (1996) and references therein]. The wide width of the melting curve is similar to that found in previous measurements of small (10−20 residues) α-helical peptides (Scholtz & Baldwin, 1992; Williams et al., 1996). Our measurements have not determined the composition of the solvated helical component nor which residues are included, although the AGH core is excluded. This portion of apoMb-I could be called molten globule-like. The IR evidence (low amide-I′ frequency, noncooperative melting, fast kinetics, and temperature independence of the kinetics) suggests that the solvated helix component is composed of a number of helices of varying length which are rapidly interconverting. It seems likely that the solvated helices are not highly localized to specific runs of polypeptide, but rather the polypeptide *between* the A and G helices contains segments that fluctuate between helical and random secondary structures. However, it is possible that the solvated helical portion of apoMb-I consists of fixed regions of helical runs of varying lengths. This type of structure would also yield the broad noncooperative melting behavior that is observed.

It might be expected that the solvated helices of the I form are stabilized by interhelical contacts since the I form of apoMb is quite compact (see the introductory portion of the paper) and at least some of these helices are almost certainly in close contact. However, the observed unfolding/folding dynamics of solvated α-helical portion(s) of the I species are very similar to an isolated helix (Williams et al., 1996). They, thus, appear to be essentially unaffected by hydrophobic (or any other) interhelical interactions, probably due to the packing density decrease of this structure compared to the native species. Such a density decrease provides ready access by water to the helices that are unfolding and, therefore, rapid replacement of intrahelix hydrogen bonds by hydrogen bonds to water. In addition, a less dense structure makes possible noncooperative displacements of the side chains (piece by piece) from the hydrophobic core. In any case, the solvated helical portion of apoMb-I behaves in the relaxation measurements just as if the helical runs were isolated from each other.

We conclude that the structure of apoMb-I is both kinetically and thermodynamically heterogeneous, consisting of the AGH core stabilized by native or native-like interhelical contacts and solvated helical portions with little or no interhelical contacts. A substantial amount of the protein also exists as a disordered conformation. This conclusion offers a reasonable structural interpretation of recent unfolding results by Jamin and Baldwin (1996), in which a burst phase was observed in the kinetics of the urea denaturation of the I form in addition to a longer time kinetics component. Such biphasic kinetics are rarely observed in the unfolding of proteins. From our results, the burst phase can now be ascribed to the melting of the solvated helical portion of apoMb-I while the long time component is due to the melting of the AGH core. These two components do not melt sequentially but rather independently, or quasi-independently, in the kinetics experiments of Jamin and Baldwin (1996), as they do in the present study (in our case, we never reach high enough temperature to melt the AGH core appreciably).

The Folding Pathway of Apomyoglobin. Combining the results from our own laboratory on apoMb-N and apoMb-I and previous studies, a quite detailed picture of the folding of apomyoglobin emerges in terms of what structures and substructures form at what times and in what manner. Since the unfolded state of any protein is very heterogeneous, it may not be fruitful to imagine specific, well-defined kinetic intermediate steps in the folding process. Rather, the process, particularly at very early times, is likely better described statistically (Dill & Chan, 1997). Nevertheless, it is possible to determine characteristic time scales. For example, we and others have shown that helical turns, or the addition of turns to an already short helix, are formed on the time scale of tens of nanoseconds, and this rate is not very temperature dependent. The unfolding time can be shorter or longer depending upon the stability of the helical segment [current results (Williams et al., 1996; Eaton et al., 1996)]. These times are consistent with a number of experiments, particularly hydrogen exchange and CD kinetics experiments [cf., Jennings and Wright (1993)], that point toward the formation of helices or nascent helical substructures at very early times in the folding pathway of apoMb.

The current results and previous studies on the I form of apoMb, taken as a model of early folding intermediate, shows that tertiary contacts already exist within the AGH core. Apart from the AGH core, it is known from our previous studies of apoMb-N that native-like tertiary contacts also form in another substructure, involving presumably the B, C, D, and E helices (Gilmanshin et al., 1997b). Indeed, the major structural difference between apoMb-N and apoMb-I is the partial gain of solvated helical substructure with the concomitant loss of native-like contacts in the latter form: both forms are heterogeneous and contain native-like helices, solvated helices, and unfolded runs of polypeptide. In neither apoMb-N nor apoMb-I have we found evidence for significant influence of interhelical interactions in the solvated

α-helical substructures. Under even more destabilizing conditions (pH 3 and low salt) than those for apoMb-I, results from both IR and fluorescence spectroscopies have established that apoMb consists mainly of the AGH core with less solvated helical substructure than in the I form (Gilmanshin et al., 1997a). In this state, which we have called apoMb-E (for extended), the AGH core is found to be characterized by a native-like IR amide-I′ (ca. 1650 cm^{-1}) frequency and also to melt cooperatively with a sigmoidal-like melting transition (paper in preparation).

The essential picture of folding of apoMb that emerges from all this is that there are specific peptide hot spots serving as nucleation points forming native-like contacts, one for AGH core formation and one for the B, C, D, and E interhelical interactions. These interactions must depend on a propensity of forming helical loop structures since the loop, at least presumably, must be present to form the observed native-like contacts. The studies of Wright and co-workers suggest that this is the case at least for the G-H helical hairpin (Shin et al., 1993). The formation of the AGH core is quite important since this event essentially ties down the two ends of the molecule with a large reduction in entropy. That the AGH core is found to contain native-like contacts in apoMb-E in the absence of other major helical structure suggests that it can be formed along the folding pathway even in the absence of hydrophobic collapse of any other part of apoMb. It may be that the number of key residues determining these substructures is not large in number. For the AGH core, the number of residues may only involve those at the intersection of the three helices.

Therefore, a great deal is known about the important structural intermediates along the folding pathway; what about the timing of events? Helix formation and/or lengthening occurs on the tens of nanoseconds time scale, as stated above and implied in the current results. Given a reasonable probability for appropriately populated helical runs, it is not known how long it takes the two ends of apoMb to come together to form the AGH core. Considerations based purely on diffusion suggest that this event can occur within microseconds [see, e.g., Karplus and Weaver (1994)], although it could proceed faster than the random diffusion limit should the peptide sequence of apoMb be coded to form strategic turns, or runs of helix that effectively shorten the loop connecting the two ends. From stopped-flow studies, it is known that this association occurs in less than a few milliseconds at ca. 5 °C [the AGH core is intact within the burst phase of stopped flow experiments (Jennings and Wright, 1993)]. The recent T-jump relaxation measurements by Ballew et al. (1996) suggest the possibility that the AGH core may form in about 5 μs at 10−15 °C. The actual timing of this specific event for apomyoglobin and its temperature dependence can be determined by relaxation experiments of the type reported here and such experiments are currently underway. The formation time of the B, C, D, and E helices association is much slower, about 280 μs at 60 °C (Gilmanshin et al., 1997b) and 1−10 s at ca. 5 °C (Jennings and Wright, 1993).

ACKNOWLEDGMENT

We thank Dr. Barry Honig for stimulating discussions.

REFERENCES

Ballew, R. M., Sabelko, J., & Gruebele, M. (1996) *Nat. Struct. Biol. 3*, 923−926.

Callender, R., Gilmanshin, R., Dyer, R. B., & Woodruff, W. (1994) *Phys. World 7*, 41−45.

Causgrove, T. P., & Dyer, R. B. (1993) *Biochemistry 32*, 11985−11991.

Chirgadze, Y. N., Shestopalov, B. V., & Venyaminov, S. Y. (1973) *Biopolymers 12*, 1337−1351.

Crumpton, M. J., & Polson, A. (1965) *J. Mol. Biol. 11*, 722−729.

Dill, K. A., & Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.

Eaton, W. A., Thompson, P. A., Chan, C.-K., Hagen, S. J., & Hofrichter, J. (1996) *Structure 4*, 1133−1138.

Eliezer, D., Jennings, P. A., Wright, P. E., Doniach, S., Hodgson, K. O., & Tsuruta, H. (1995) *Science 270*, 487−488.

Evans, P. A., & Radford, S. E. (1994) *Curr. Opin. Struct. Biol. 4*, 100−106.

Evans, S. V., & Brayer, G. D. (1988) *J. Biol. Chem. 263*, 4263−4268.

Fink, A. L., Calciano, L. J., Goto, Y., & Palleros, D. (1991) in *Conformstions and forces in protein folding* (Nall, B. T., & Dill, K. A., Eds.) pp 169−174, American Association for the Advancement of Science, Washington, DC.

Gast, K., Damaschun, H., Misselwitz, R., Müller-Frohne, M., Zirwer, D., & Damaschun, G. (1994) *Eur. Biophys. J. 23*, 297−305.

Gilmanshin, R., Dyer, R. B., Woodruff, W. H., & Callender, R. H. (1997a) *Protein Sci. 6*, 2134−2142.

Gilmanshin, R., Williams, S., Callender, R. H., Woodruff, W. H., & Dyer, R. B. (1997b) *Proc. Natl. Acad. Sci. U.S.A. 94*, 3709−3713.

Griko, Y. V., Privalov, P. L., Venyaminov, S. Y., & Kutyshenko, V. P. (1988) *J. Mol. Biol. 202*, 127−138.

Gruenewald, B., Nicola, C. U., Lustig, A., Schwarz, G., & Klump, H. (1979) *Biophys. Chem. 9*, 137−147.

Haris, P. I., & Chapman, D. (1995) *Biopolymers 37*, 251−263.

Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544−1548.

Jackson, M., & Mantsch, H. H. (1995) *Crit. Rev. Biochem. Mol. Biol. 30*, 95−120.

Jamin, M., & Baldwin, R. L. (1996) *Nat. Struct. Biol. 3*, 613−618.

Jennings, P. A., & Wright, P. E. (1993) *Science 262*, 892−896.

Johnson, R. S., & Walsh, K. A. (1994) *Protein Sci. 3*, 2411−2418.

Karplus, M., & Weaver, D. L. (1994) *Protein Sci. 3*, 650−668.

Kay, M. S., & Baldwin, R. L. (1996) *Nat. Struct. Biol. 3*, 439−445.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.

Lecomte, J. T. J., Kao, Y. H., & Cocco, M. J. (1996) *Proteins: Struct., Funct., Genet. 25*, 267−285.

Ptitsyn, O. B. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 243−300, W. H. Freeman & Co, New York.

Reinstädler, D., Fabian, H., Backmann, J., & Naumann, D. (1996) *Biochemistry 35*, 15822−15830.

Reisdorf, W. C., Jr., & Krimm, S. (1996) *Biochemistry 35*, 1383−1386.

Rothgeb, T. M., & Gurd, F. R. N. (1978) *Methods Enzymol. 52*, 473−486.

Schmid, F. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 197−241, W. H. Freeman & Co, New York.

Scholtz, J. M., & Baldwin, R. L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 95−118.

Schwarz, G. (1965) *J. Mol. Biol. 11*, 64−77.

Shin, H.-C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6356−6364.

Susi, H., & Byler, D. M. (1986) *Methods Enzymol. 130*, 290−311.

Teale, F. W. J. (1959) *Biochim. Biophys. Acta 35*, 543.

Uversky, V. N., & Ptitsyn, O. B. (1996) *Folding Des. 1*, 117−122.

Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., & Dyer, R. B. (1996) *Biochemistry 35*, 691−697.

Zwanzig, R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9801−9804.

BI970634R