Structure and Dynamics of the Acid-Denatured Molten Globule State of α-Lactalbumin: A Two-Dimensional NMR Study†

Andrei T. Alexandrescu,‡ Philip A. Evans,§ Maureen Pitkeathly,‡ Jean Baum,∥ and Christopher M. Dobson*,‡

Oxford Centre for Molecular Sciences and Inorganic Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QR, England, Department of Biochemistry, Cambridge University, Tennis Court Road, Cambridge CB2 1QW, England, and Department of Chemistry, Rutgers University, Piscataway, New Jersey 08854

Received July 21, 1992; Revised Manuscript Received November 2, 1992

ABSTRACT: Two-dimensional ^{1}H-NMR spectroscopy has been used to study the acid-denatured molten globule (A-state) of α-lactalbumin. The NMR spectra show that chemical shift dispersion is limited but significantly greater than that expected for a random coil conformation. The small chemical shift dispersion of side-chain resonances in the A-state together with line broadening associated with conformational averaging indicates that most of the long-range tertiary structure in the A-state is likely to be nonspecific. Side-chain resonances in the A-state are generally shifted somewhat upfield of random coil values; this and the observation of a large number of interresidue NOEs, however, indicate that some side-chain interactions, at least at the level of hydrophobic clustering, exist in the A-state. Analysis of NOESY spectra shows no evidence for an ordered structure for either of the two major clusters of aromatic residues which in the native structure make up part of the hydrophobic core of the helical domain of the native protein. A new aromatic cluster in the A-state which results from rearrangement of the side chains of Tyr103, Trp104, and His107 from their native state positions was, however, detected by a number of well-defined interresidue NOE effects. Similar NOE patterns are observed in a peptide corresponding to residues 101–110 of α-lactalbumin in trifluoroethanol, suggesting that the nonnative structure in the 101–110 region of the A-state is not dependent on specific interactions with the rest of the chain. Trapping experiments indicate that amide protons from regions of the sequence which in the native state are helical are among those strongly protected from solvent exchange in the A-state; those from one of the helices (the C helix) were specifically identified. Taken together, these results reinforce a model of the A-state which has stable regions of localized secondary structure but a largely disordered tertiary structure.

The mechanism by which an unstructured polypeptide chain folds to its final three-dimensional structure remains one of the central puzzles in biochemistry. The short lifetimes of kinetic intermediates in the folding reaction mean that direct methods for the detailed structural characterization of these forms are at present limited. Nevertheless, it is becoming increasingly evident that structural information is needed in order to formulate mechanisms for protein folding. One strategy to obtain such information is to study partially folded states of proteins stable at equilibrium (Kuwajima, 1989; Kim & Baldwin, 1990; Dobson, 1991, 1992). The study of states with intermediate levels of order is of interest in its own right as it will aid the understanding of how native structure, or for that matter structure in proteins under nonnative conditions, is organized. Residual structure in equilibrium intermediates is also often related to that found in kinetic intermediates. Of particular interest in this regard are molten globules, which are partially folded states that show native-like secondary structure characteristics but little evidence for specific tertiary structure.

The molten globule state of α-lactalbumin is a well-defined equilibrium intermediate that shares many properties of a kinetic intermediate on the folding pathway (Kuwajima et al., 1985; Ikeguchi et al., 1986). It is possible partially to unfold α-lactalbumin to a molten globule state under a variety of denaturing conditions (e.g., extremes of pH, removal of the bound Ca^{2+} ion, reduction of disulfide bonds, high temperature). A number of techniques suggest that the residual structure obtained under these varied conditions is very similar (Kuwajima, 1989). The most extensively characterized molten globule is the one obtained by acid-induced unfolding of α-lactalbumin at pH 2.0 (the A-state). Intrinsic viscosity measurements (Dolgikh et al., 1985), quasielastic light scattering (Gast et al., 1986), and high-angle diffuse X-ray scattering (Damaschun et al., 1986) indicate that the A-state of α-lactalbumin is nearly as compact as the native state. Far UV[1] CD (Kuwajima, 1977; Dolgikh et al., 1981) suggests that the amount of secondary structure in the A-state is comparable to that found in the native state. By contrast, the absence of a significant near-UV signal in CD spectra (Dolgikh et al., 1981), increased binding of the hydrophobic probe ANS (Mulqueen & Kronman, 1982; Semisotnov et al., 1991), and

† This work was supported by the U.K. Science and Engineering and Medical Research Councils. Portions of this work were supported by NATO-NSF Grant RCD-8953799 awarded to A.T.A. in 1990, and by NIH Grant GM-45302 awarded to J.B. P.A.E. is a member of the Cambridge Centre for Molecular Recognition. C.M.D. is a member of the Oxford Centre for Molecular Sciences.

* To whom correspondence should be addressed.
‡ Oxford University.
§ Cambridge University.
∥ Rutgers University.

[1] Abbreviations: BLA, bovine α-lactalbumin; HLA, human α-lactalbumin; GPLA, guinea pig α-lactalbumin; HEWL, hen egg white lysozyme; TFE, trifluoroethanol; ANS, 1-anilinonaphthalene-8-sulfonate; TSP, 3-(trimethylsilyl)propionate-d_4; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; DQF-COSY, double-quantum filtered two-dimensional *J*-correlated spectroscopy; NOESY, two-dimensional NOE spectroscopy; ROESY, two-dimensional spin-locked (rotating frame) NOE spectroscopy; CD, circular dichroism; UV, ultraviolet. Amino acid residues are designated by their standard one-letter abbreviations. The nomenclature used to identify specific protons is that of the Brookhaven Protein Data Bank (PDB).

0006-2960/93/0432-1707$04.00/0

disulfide rearrangement experiments (Ewbank & Creighton, 1991) indicate that the tertiary structure of the A-state is disordered compared to that of the native state. These techniques, however, give only a global picture of the A-state. Important issues regarding the residual structure in the A-state remain unresolved. In particular, which segments of the chain are ordered in the A-state? How similar is the residual structure to that of the native state? How labile is the residual structure? The only methods currently capable of providing structural information for proteins at atomic resolution are X-ray diffraction and NMR spectroscopy. The usefulness of NMR spectroscopy for structural studies of denatured states has already been demonstrated (Baum et al., 1989; Hughson et al., 1990; Harding et al., 1991; Jeng & Englander, 1991; Evans et al., 1991; Neri et al., 1992). By using amide trapping techniques in conjunction with NMR spectroscopy of the native state of GPLA, it has been possible to show that regions of the sequence that correspond to two helices (B and C) in the native state give rise to persistent secondary structure in the A-state of GPLA (Baum et al., 1989; Chyan et al., 1993). 1D NMR experiments of the A-state of α-lactalbumin (Dolgikh, et al., 1985; Baum et al., 1989), as well as the molten globule states of a number of other proteins [e.g., Jeng and Englander (1991)], have suggested that there is a high degree of disorder of the side chains of at least the majority of residues in molten globule states. This disorder is manifested in NMR spectra as increased line broadening and decreased chemical shift dispersion of resonances. In this paper we examine the utility of 2D NMR spectroscopy (COSY, NOESY) for direct structural characterization of the A-state of α-lactalbumin. In principle, multidimensional NMR methods can provide structural detail at atomic resolution and complement studies of the global properties of the A-state. Furthermore, because these methods are not limited by the solvent exchange properties of the native state, it is possible to obtain information about tertiary interactions and nonnative structure in the A-state which complements the information obtained from the amide trapping technique.

MATERIALS AND METHODS

Materials. D_2O (99.96% D, 99.996% D), DCl (99% D), and NaOD (99% D) were obtained from Sigma; CF_3CD_2OD (99% D) was obtained from CEN Saclay (France). BLA and HLA were obtained from the Sigma Chemical Co.; GPLA was prepared according to a previously published method (Baum et al., 1989). Protein preparations were checked by native capillary electrophoresis in 100 mM phosphoric acid buffer and/or SDS–PAGE and were in excess of 90% pure. These were dialyzed against an excess of $CaCl_2$ (0.1 M), and a further three times against water at pH 7.5 before use. Reduced carboxymethylated BLA (8CM-BLA) was also obtained from Sigma. Capillary electrophoresis indicated the presence of three closely spaced bands which made up 69%, 18%, and 10% of the sample. Amino acid analysis indicated the presence of approximately 7.5 mol of carboxymethylated cysteine per mole of protein and only trace amounts of free cysteine. The heterogeneity observed for 8CM-BLA by native capillary electrophoresis may be due to rapid carboxymethylation of the unique methionine at position 90 (Castellino & Hill, 1970) or to incomplete carboxymethylation of the eight cysteine residues. No heterogeneity was detected in NMR spectra of 8CM-BLA.

Synthesis and Purification of a Peptide Corresponding to Residues 101–110 of HLA (H101–110). The sequence of HLA (I101-D102-Y103-W104-L105-A106-H107-K108-A109-L110) was chosen for the H101–110 peptide model because a crystal structure is available for the native state of HLA; the corresponding region of this sequence in BLA differs only by the substitution of asparagine for aspartate at position 102. In GPLA, the alanine at position 109 is replaced by a proline. Peptide synthesis was carried out on an Applied Biosystems model 430A automated peptide synthesizer using standard *N*-(9-fluorenyl)methoxycarbonyl chemistry procedures modified to cap unreacted chains. Treatment of the protected resin-bound peptide with aqueous trifluoroacetic acid in the presence of ethanedithiol and thioanisole afforded the crude peptide which was purified by reverse-phase HPLC. Amino acid analysis was obtained by using an Applied Biosystems 420A analyzer after precolumn derivatization with phenyl isocyanate. The amino acid composition A(2.0), D/N(1.2), H(1.1), I(0.9), K(1.2), L(2.0), Y(0.9), W(not determined) is consistent with the expected sequence for the H101–110 peptide.

NMR Spectroscopy. TSP was used as an internal chemical shift standard. The pH values of samples were adjusted by addition of 0.4–2.0-μL aliquots of NaOD or DCl. All pH values were measured at room temperature with a (Philips CE2/180) glass electrode; pH values reported in this work are apparent pH values and were not corrected for the deuterium isotope effect or for the effect of CF_3CD_2OD.

NMR spectra were recorded using Bruker AM 500 and 600 spectrometers of the Oxford Centre for Molecular Sciences and processed using DISNMR software. Relaxation delays of 1.0–1.5 s were left between pulse cycles for NMR data collected on proteins; delays of 2.0–3.0 s were used for work on the H101–110 peptide. A presaturation pulse was used to suppress the residual water peak. For 2D NMR spectroscopy of the A-state, we were unable to obtain NMR spectra of sufficient quality for detailed interpretation at temperatures below 315 K. It should be noted, therefore, that the A-state described in this work may be less structured than A-states studied at lower temperatures.

Assignment of the C-Helix in the Native State of BLA. Sequence-specific ^{1}H-NMR assignments (Table II) were obtained by using data from 2D DQF-COSY (Rance et al., 1983), NOESY (Anil Kumar et al., 1980), RELAY (Bax & Drobny, 1985), and double-RELAY experiments. The NH protons of the C-helix of BLA were initially identified by comparing an NH–NH sequential walk in NOESY spectra of BLA to that reported for GPLA (Baum et al., 1989). The identity of the sequence was confirmed by obtaining spin-system assignments for residues in the C-helix by using COSY, RELAY, and double-RELAY experiments.

Amide Trapping Hydrogen Exchange Experiments. Hydrogen exchange experiments were performed at a temperature of 298 K. Samples were initially dissolved at neutral pH in 99.996% D_2O. For the native state experiments, the pH was adjusted to 5.3, and NMR spectra were collected as a function of time. For the amide trapping experiments (Baum et al., 1989), the pH was adjusted to 2.0. After incubation at pH 2.0 in D_2O for various lengths of time, exchange in the MG state was quenched by adjusting the pH to 5.0 (native state). Samples were subsequently lyophilized and taken up in 99.996% D_2O, pH 5.0, for NMR spectroscopy.

Rates for amide proton exchange in the native and the A-state were obtained from the decay of intensity (peak height) of amide proton resonances in 1D NMR spectra with time. The (1D) NMR spectra were phased and baseline corrected in the amide region, and the intensities of the amide proton

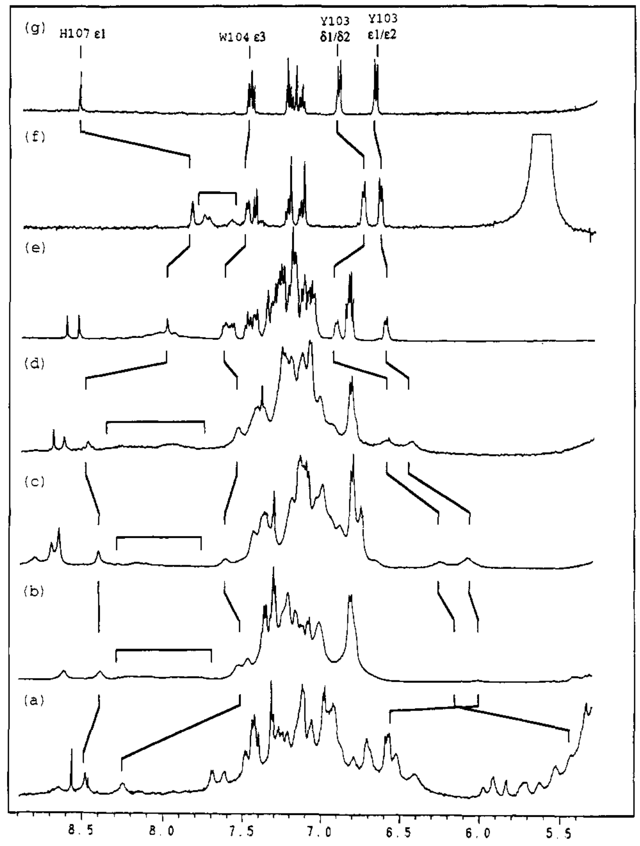

FIGURE 1: 1D NMR spectra of α-lactalbumin and the H101–110 peptide. (a) Native BLA: 2.0 mM protein in D_2O, pH 4.2, 315 K. (b) A-state HLA: 2.1 mM protein in D_2O, pH 2.0, 315 K. (c) A-state GPLA: 2.1 mM protein in D_2O, pH 2.0, 315 K. (d) A-state BLA: 2.3 mM protein in D_2O, pH 2.0, 315 K. (e) A-state BLA in the presence of 50% (v/v) CF_3CD_2OD/50% D_2O: 1.0 mM protein, pH 2.0, 315 K. (f) H101-110 peptide in 95% (v/v) CF_3CD_2OD/5% D_2O: 1.0 mM peptide, pH 2.1, 279 K. (g) H101–110 peptide in D_2O: 1.0 mM peptide, pH 2.0, 279 K. Resonances of Y103, W104, and H107 are indicated in the various forms. The resonances indicated by brackets are due to unexchanged amide protons.

resonances were calibrated relative to the intensities of two or more resonances from nonexchangeable protons.

Assignment of the H101–110 Peptide. Spin system assignments for the H101–110 peptide were obtained from 2D NOESY and COSY spectra using standard procedures (Wüthrich, 1986). Sequence-specific assignments were obtained from sequential NH_i–NH_{i+1} and αCH_i–NH_{i+1} NOE connectivities (Wüthrich, 1986). Full details of these assignments will be given elsewhere (A.T.A. and C.M.D., manuscript in preparation).

RESULTS

General Features of the NMR Spectra. Figure 1 shows 1D NMR spectra of native BLA, the A-states of BLA, HLA, and GPLA, the A-state of BLA in the presence of 50% (v/v) TFE, and the H101–110 peptide in 95% TFE and in water. The chemical shift dispersion for the A-states is diminished considerably compared to that for the native states, but for many resonances line broadening is increased. Of the three α-lactalbumin variants, BLA gave on average the sharpest lines in the A-state and was thus the most amenable to 2D NMR spectroscopy. Line widths were found to be independent of protein concentration in the range 0.1–4.0 mM, and equilibrium sedimentation measurements of the A-state of BLA (Dolgikh et al., 1985) indicate that the major species in the concentration range 0.1–1.3 mM at low ionic strength is a monomer. Protein association, therefore, is unlikely to be a major factor contributing to line broadening in the NMR spectra of the A-state. Resonance broadening of the type observed for the A-state of BLA has been observed for the molten globule states of a number of proteins including GPLA (Baum et al., 1989), cytochrome *c* (Jeng & Englander, 1991), and the F2 fragment of *Escherichia coli* tryptophan synthase β2-subunit (Chaffote et al., 1991) and has been attributed to conformational averaging on the millisecond time scale. NMR spectra of the native states of the three α-lactalbumins are similar to each other (Alexandrescu et al., 1992). The NMR spectra of the three proteins in their A-states also show close similarity suggesting that the essential features of the A-state are conserved in the three α-lactalbumins.

Distribution of Chemical Shifts. Figure 2 shows a 2D DQF-COSY spectrum of the A-state of BLA. It is possible to identify regions in this spectrum which contain resonances arising from distinct types of protons. The chemical shift parameters of protons resonating in these regions are compared to those expected for a random coil conformation in Table I. Of particular interest are the aromatic ring protons (Figure 2b) which occur in a relatively isolated region of the spectrum and have been extensively characterized in the native state of BLA (Alexandrescu et al., 1992). In an idealized random coil conformation the aromatic residues of BLA [three histidines (H32, H68, H107), four tyrosines (Y18, Y36, Y50, Y103), four phenylalanines (F9, F31, F53, F80), four tryptophans (W26, W60, W104, W118)] would be expected to give rise to only four spin systems, one corresponding to each different type of aromatic side chain. The four phenylalanine spin systems (7.15, 7.21–7.27, and 7.30 ppm) show little chemical shift dispersion (Figure 2b). Of the remaining aromatic spin systems, however, it is possible to distinguish three of the four tyrosines, all of the three histidines, and all of the four tryptophans. More generally, the groups of protons identified in Table I almost all show dispersion greater than that expected for a random coil conformation. The chemical shift dispersion for α-lactalbumin in the presence of 9 M urea (Baum et al., 1989), or in denatured HEWL (a protein with around 40% sequence homology to α-lactalbumin), is substantially smaller than in the A-state of BLA (Evans et al., 1991). The amount of residual structure in denatured HEWL has been shown by a variety of experimental observations to be much lower than in the A-state of α-lactalbumin (Kuwajima, 1989). This, therefore, strongly suggests that the shifts observed in the A-state of BLA reflect a more structured denatured state for α-lactalbumin than for lysozyme.

It is possible to identify some trends in the direction of chemical shift deviations from random coil values (Table I). The αCH resonances are distributed mainly upfield of random coil values, which is suggestive (Wishart et al., 1991) of the presence of α-helical secondary structure in the A-state. More specifically, groups E and D in Figure 2 each manifest about 10 αCH resonances which are shifted upfield of random coil values. The lack of many αCH resonances shifted downfield of random coil values suggests (Wishart et al., 1991) that well-defined β-sheet structure is not present to a large extent in the A-state. Furthermore, strong αCH–αCH NOEs which are associated with antiparallel β-sheet structure (Wüthrich, 1986) are almost entirely absent in NOESY spectra of the A-state. In the native-state X-ray structure, only six residues were reported to be involved in well-defined β-sheet structure

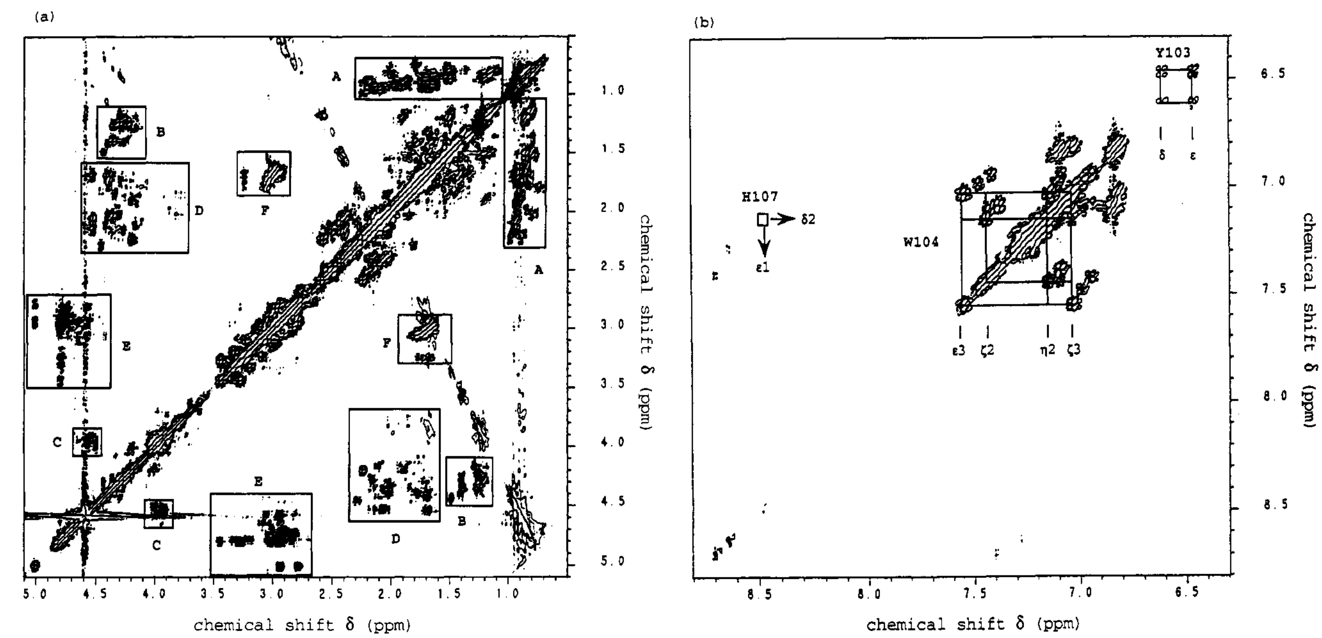

FIGURE 2: DQF-COSY spectrum of BLA in the A-state. (a) Aliphatic region; (b) aromatic region. Conditions: 4.0 mM protein in D_2O, pH 2.0, 320 K. The data set consisted of 1024 t_1 increments of 2K data points. The data were zero-filled once in ω_2 and twice in ω_1 and multiplied by a $\pi/15$-shifted sine bell function in both dimensions prior to the 2D transform. The following coupling regions are indicated in panel a: (A) V (βCH–γCH_3), I (βCH–γCH_3, γCH_2–δCH_3), L (γCH–δCH_3); (B) A (αCH–βCH_3), T (βCH–γCH_3); (C) S (αCH–βCH_2); (D) E, I, K, L, M, P, Q, R, V (αCH–βCH_x); (E) C, D, F, H, N, W, Y (αCH–βCH_2); (F) K (ϵCH_2–δCH_2), R (δCH_2–γCH_2). The aromatic proton spin systems of Y103, W104, and H107 are indicated in panel b. The H107 ϵ1CH–δ2CH cross-peak (box) could only be observed at lower contour levels. Chemical shift parameters for the groups of protons identified in this figure are tabulated in Table I.

Table I: Chemical Shift Dispersion in Acid-Denatured α-Lactalbumin[a]

resonances	A-state (ppm)	random coil (ppm)	spectral region
NH (all)	8.88–7.45	8.75–8.09	Figure 3b
A (α-CH)	4.43–4.17	4.35	Figure 2a, B
A (β-CH_3)	1.49–1.30	1.39	Figure 2a, B
V, I, L (γ,δ-CH_3)	1.00–0.71	0.97–0.89	Figure 2a, A
T (γ-CH_3)	1.40–1.18	1.23	Figure 2a, B
T (β-CH)	4.40–4.23	4.35	Figure 2a, B
S (α-CH)	4.58–4.49	4.50	Figure 2a, C
S (β-CH)	4.03–3.89	3.88	Figure 2a, C
F, Y, H, W, C, D, N			
α-CH	4.86–4.44[b]	4.76–4.60	Figure 2a, E
β-CH	3.45–2.69	3.32–2.75	Figure 2a, E
I, L, K, R, V, E, Q, M, P			
α-CH	4.53–3.76	4.52–4.18	Figure 2a, D
β-CH	2.28–1.67	2.28–1.65	Figure 2a, D
aromatics			
W (δ1)	7.19–7.17	7.24	Figure 3b
W (ϵ1)	10.04–9.95	9.98	Figure 3b
W (ϵ3)	7.47–7.41	7.65	Figure 2b
W (ζ3)	6.99–6.94	7.17	Figure 2b
W (η2)	7.12–7.08	7.24	Figure 2b
W (ζ2)	7.42–7.37	7.50	Figure 2b
Y (ϵ)	6.85–6.82	6.86	Figure 2b
Y (δ)	7.10–7.03	7.15	Figure 2b
H (ϵ1)	8.69–8.62	8.70	Figure 2b
H (δ2)	7.40–7.28	7.40	Figure 2b
F (ϵ)	7.30	7.39	Figure 2b
F (δ)	7.15	7.30	Figure 2b
F (ζ)	7.27–7.21	7.34	Figure 2b

[a] Resonances from Y103, W104, and H107 are excluded from this analysis (see Table III). Random coil values are from Wüthrich (1986) and Evans et al. (1991). [b] In this group, an additional two unidentified αCH resonances occur at 5.02 and 5.38 ppm.

(Acharya et al., 1989); the NMR data therefore suggest that the A-state does not posses significant amounts of persistent native or nonnative β-sheet. The resonances corresponding to valine, isoleucine, and leucine methyl groups, as well as those of aromatic ring protons, are generally distributed upfield of random coil values. This trend of side-chain resonances to resonate upfield of random coil values in a denatured protein has been noted previously, though to a lesser extent, for denatured HEWL and has been attributed to averaged ring-current effects associated with nonspecific hydrophobic clustering of side chains (Evans et al., 1991).

Evidence for Secondary Structure. Figure 3 compares the amide to aliphatic regions of NOESY spectra of the native and A-states of BLA. The greater chemical shift dispersion and the greater number of cross-peaks in the former are readily apparent. It is notable, however, that the A-state of BLA also manifests a large number of NOE effects. Some of these, such as the NH–NH cross-peaks and the aromatic–methyl cross-peaks (see also Figure 7a) must correspond to at least i – (i+1) interresidue interactions. A group of unresolved resonances which correspond to the $\epsilon 1/\epsilon 2$CH ring protons of Y18, Y36, and Y50 exhibit a number of NOE effects to the methyl region of the spectrum (e.g., to isoleucine, valine, or leucine residues). An analysis of the sequence of BLA indicates that, at a minimum, these NOEs correspond to i – (i+2) interresidue interactions. NOE effects can arise either from protons constrained in proximity by a rigid structure or from transient conformations if the lifetimes of such conformations are long enough to allow cross-relaxation effects to build up (Fejzo et al., 1991). Although it is unlikely that the A-state can be described as a unique conformation, the large number of interresidue NOEs, in particular strong NOEs between side chains, indicate that distinct conformational preferences exist.

The NH–NH region of the NOE spectrum is of interest in light of our investigation of secondary structure in the A-state by amide proton solvent exchange methods. Direct measurement of amide proton bulk exchange rates shows that about 20 amide protons are significantly protected in the A-state of GPLA (Baum et al., 1989; Chyan et al., 1993). Similar numbers of amide protons show significant protection in the A-states of BLA and HLA (J.B. and C.M.D., unpublished results). In the case of BLA, however, solvent

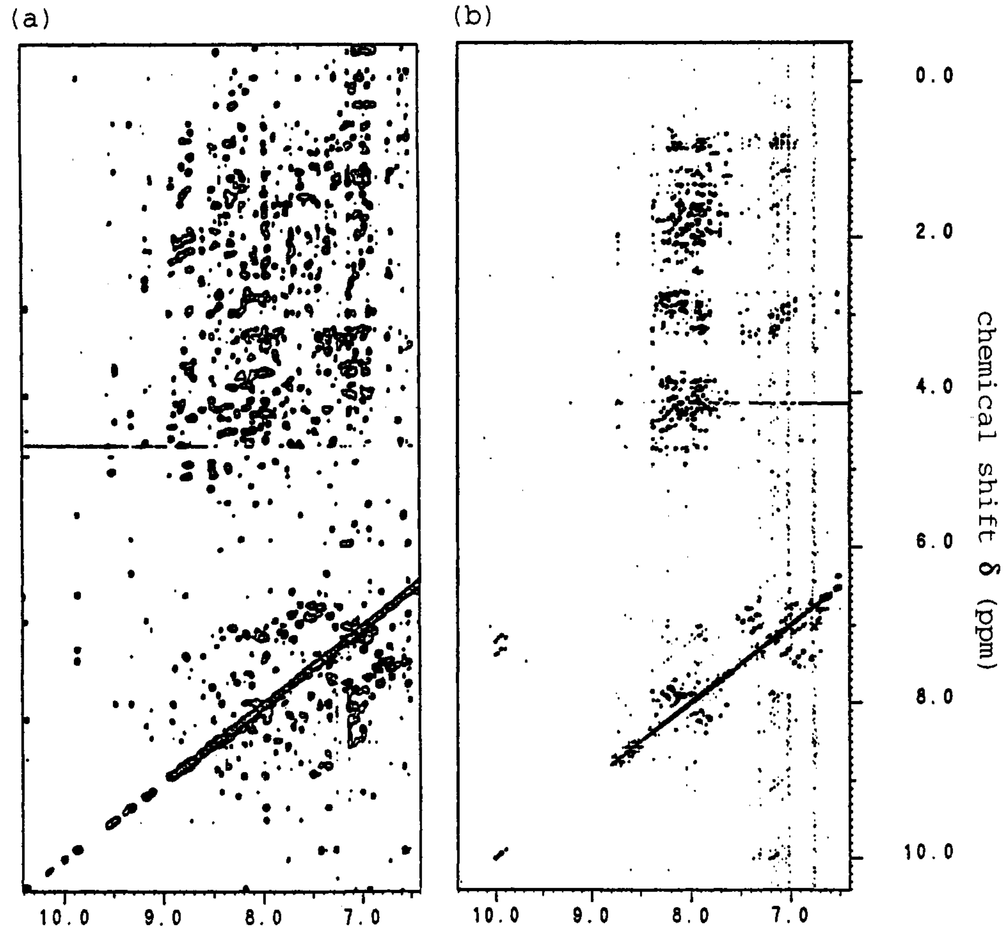

FIGURE 3: NOESY spectra of BLA in (a) the native state and (b) the A-state. Conditions: (a) 7.0 mM protein in 90% H_2O/10% D_2O, pH 5.8, 310 K, τ_m = 150 ms. The data set consisted of 512 t_1 increments of 2K data points. The data were zero-filled twice in ω_1 and multiplied by a $\pi/8$-shifted square sine bell function in both dimensions prior to the 2D transform. (b) 3.0 mM protein in 90% H_2O/10% D_2O, pH 2.0, 315 K, τ_m = 200 ms. The data set consisted of 1000 t_1 increments of 2K data points. The data were zero-filled once in ω_1 and multiplied by a $\pi/10$-shifted square sine bell function in both dimensions prior to the 2D transform. In the spectrum of the A-state (b), the four pairs of cross-peaks at ω_2 10.0 ppm and ω_1 7.2 and 7.4 ppm are due to intraresidue NOE effects between each of the tryptophan indole Nε1 protons and the corresponding Cδ1 and Nζ2 ring protons, respectively.

exchange in the A-state is somewhat faster than in GPLA and HLA. That the number of protected amide protons is similar in the three proteins suggests that their A-states share similar structural properties. In the case of GPLA, amide trapping experiments (Baum et al., 1989) indicate that regions of the sequence corresponding to the B and C helices in the native state (Figure 4a) account for almost all of the significantly protected amide protons in the A-state (Baum et al., 1989; Chyan et al., 1993). We have been able to confirm that the amides corresponding to the C-helix in the native structure are among those most strongly protected from solvent exchange in both the native and A-states of BLA (Figure 5, Table II). In addition to amide protons from the C-helix, a number of additional amide protons are protected from exchange in the A-state of BLA. We have not yet, however, been able to obtain native-state assignments for these. In particular, we have not been able to identify the B-helix amide protons in the native state of BLA. The indole side chain proton of W26 which is in the B-helix, however, is protected in the A-states of both BLA (Figure 5) and GPLA (Chyan et al., 1993). The exchange rates for C-helix amide protons in the A-state of BLA are somewhat faster than in the A-state of GPLA. Thus the exchange rate of the NH proton of L96 in the A-state of BLA is 6.2×10^{-5} s^{-1} at pH 2.0 and 298 K, which is a factor of more than 70 slower than the exchange rate predicted for a leucine residue in an unstructured conformation (Jeng & Englander, 1991). The corresponding "protection factor" for the NH proton of L96 in the A-state of GPLA is 100 (Chyan et al., 1992). The C-helix, however, is much more stable in the native state than in the A-state: the exchange rate of the NH proton of L96 in the native state of BLA is 1.8×10^{-8} s^{-1} at pH 5.3 and 298 K, corresponding to a protection factor of 8×10^6.

Figure 3b reveals about 20–30 identifiable NH–NH cross-peaks in the NOESY spectrum of the A-state of BLA; this number corresponds roughly to the number of slowly exchanging protons in the A-state of BLA. The NH–NH connectivities link stretches of residues; one stretch of five residues and two stretches of eight residues could be tentatively identified. The sequence-specific assignment of these NOEs is complicated by line broadening of amide protons, which is more severe than for other types of protons in the spectrum, possibly as a result of heteronuclear N–H dipolar coupling. The low resolution of the COSY and HOHAHA spectra of the A-state precluded the identification of enough unique side-chain spin systems to assign stretches of NH NOEs unambiguously. Other distinguishing features of α-helical structure such as small NH_i–αCH_{i+1} NOEs and small $^3J_{CH-NH}$ coupling constants could also not be diagnosed because of severe

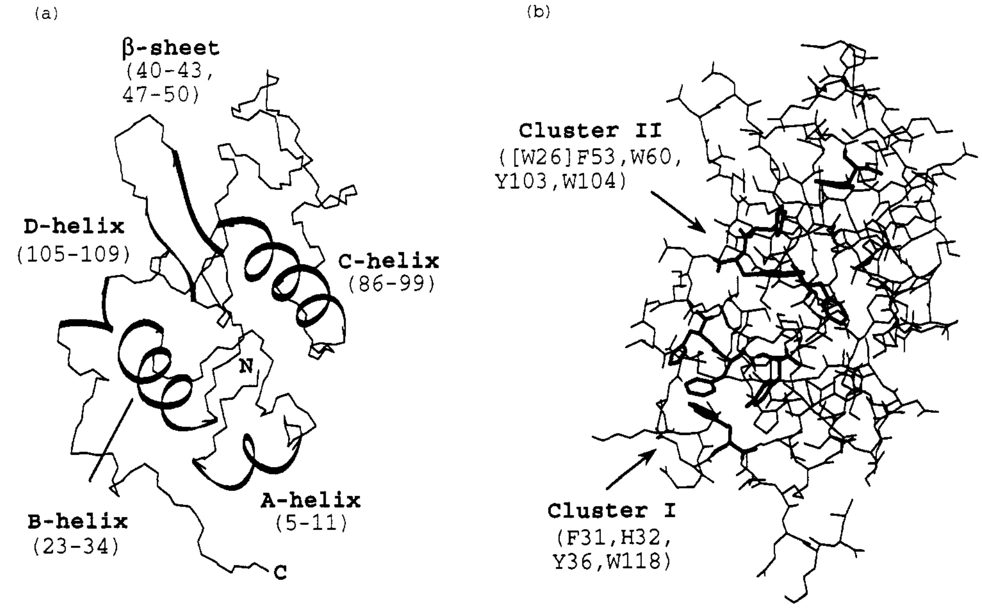

FIGURE 4: Views of the X-ray structure of HLA (Acharya et al., 1991) showing (a) backbone Cα, C′, and N atoms and (b) all atoms. In panel a, regular elements of secondary structure are indicated by ribbons. In panel b, the aromatic residues which make up clusters I and II are indicated by thick lines. In BLA an additional tryptophan at position 26 is part of cluster II (Alexandrescu et al., 1992). The molecule is shown in the same orientation in panels a and b.

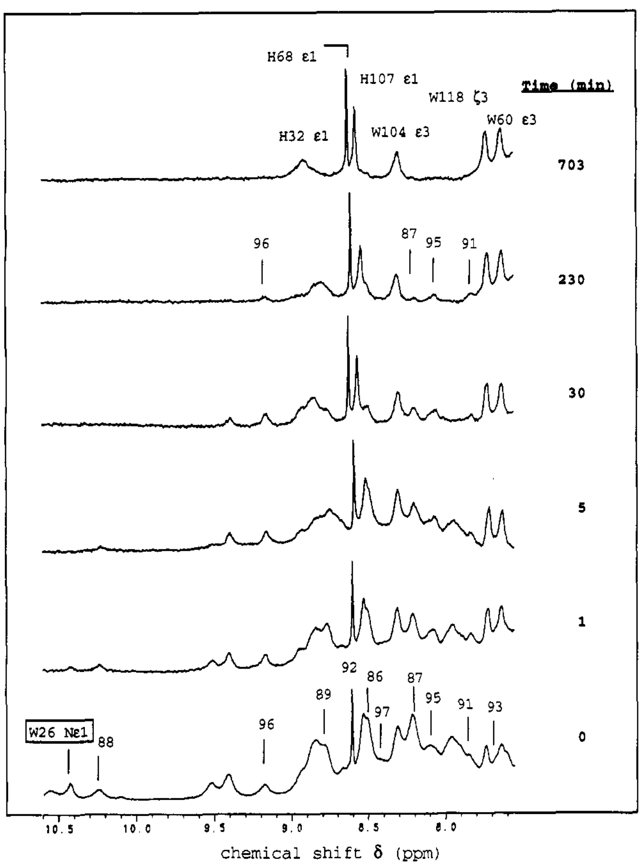

FIGURE 5: Selected time points from the trapping experiments used to probe amide exchange in the A-state of BLA. Protein concentrations were 1.3 mM. Spectra were recorded at pH 5.0 after incubation in D_2O, pH 2.0, at 298 K for the times given in the figure. The zero time point indicates the position of the C-helix amide protons and of the W26 indole NH proton. C-helix protons remaining after 230 min are indicated in the respective trace, the resonances indicated in the 703-min time point are due to nonlabile aromatic protons.

crowding in the αCH–NH region of the NOESY spectrum and because only a few NH–αCH cross-peaks (15 of the 123) could be detected in the fingerprint region of COSY spectra

Table II: Chemical Shift Values (ppm) and Assignments for C-Helix Residues in Native BLA at pH 5.83 and a Temperature of 315 K

residue	NH	αCH	βCH	γCH and others
Thr86	8.51	3.75	4.12	γCH_3 1.28
Asp87	8.19	4.43	3.40, 1.96	
Asp88	10.35	3.96		
Ile89	8.70	2.85	1.90	γCH 1.77, 1.40; γCH_3 0.53; δCH_3 0.78
Met90	7.53	4.17	2.07, 1.91	
Cys91	7.88	4.21		
Val92	8.60	4.13	1.55	γCH3 0.94, 0.86
Lys93	7.68	4.16		
Lys94	7.12	3.93		
Ile95	8.05	3.08	1.59	γCH 0.61, −2.40; γCH_3 0.42; δCH_3 −0.54
Leu96	9.13	3.84	1.51	
Asp97	8.41	3.18	2.75, 2.47	

recorded in H_2O. In native proteins α-helical structure is characterized by strong NH_i–NH_{i+1} NOEs and additional weak NH_i–NH_{i+3} NOEs. About half of the NH–NH NOEs in Figure 3b have intensities comparable to those connecting neighboring aromatic ring protons on the same residue [e.g., tryptophan ε1CH–δ1CH (2.6 Å) and ε1CH–δ1CH (2.7 Å)] and must correspond to short NH–NH distances (2.6–2.8 Å) which are characteristic of either α-helical structure or turns. The fact that NH–NH cross-peaks appear to connect stretches of residues strongly suggests that the NH–NH connectivities correspond to helical structure. At the same time, however, we did not detect any NH_i–NH_{i+3} NOEs for the stretches of NH–NH connectivities identified in the NOESY spectrum of the A-state. The absence of NH_i–NH_{i+3} NOEs may be due to the lower persistence of helical structure in the A-state, as well as to the effects of line broadening discussed above.

Properties of Aromatic Residues. In the X-ray structure of native HLA (Acharya et al., 1989, 1991), two clusters of aromatic residues are evident (Figure 4b). The first of these is made up of F31, H32, Y36, and W118. The second consists of W26, F53, W60, W104, and Y103. The presence of these clusters in solution has been confirmed by NOE spectroscopy of HLA, GPLA, and BLA in their native states (Alexandrescu

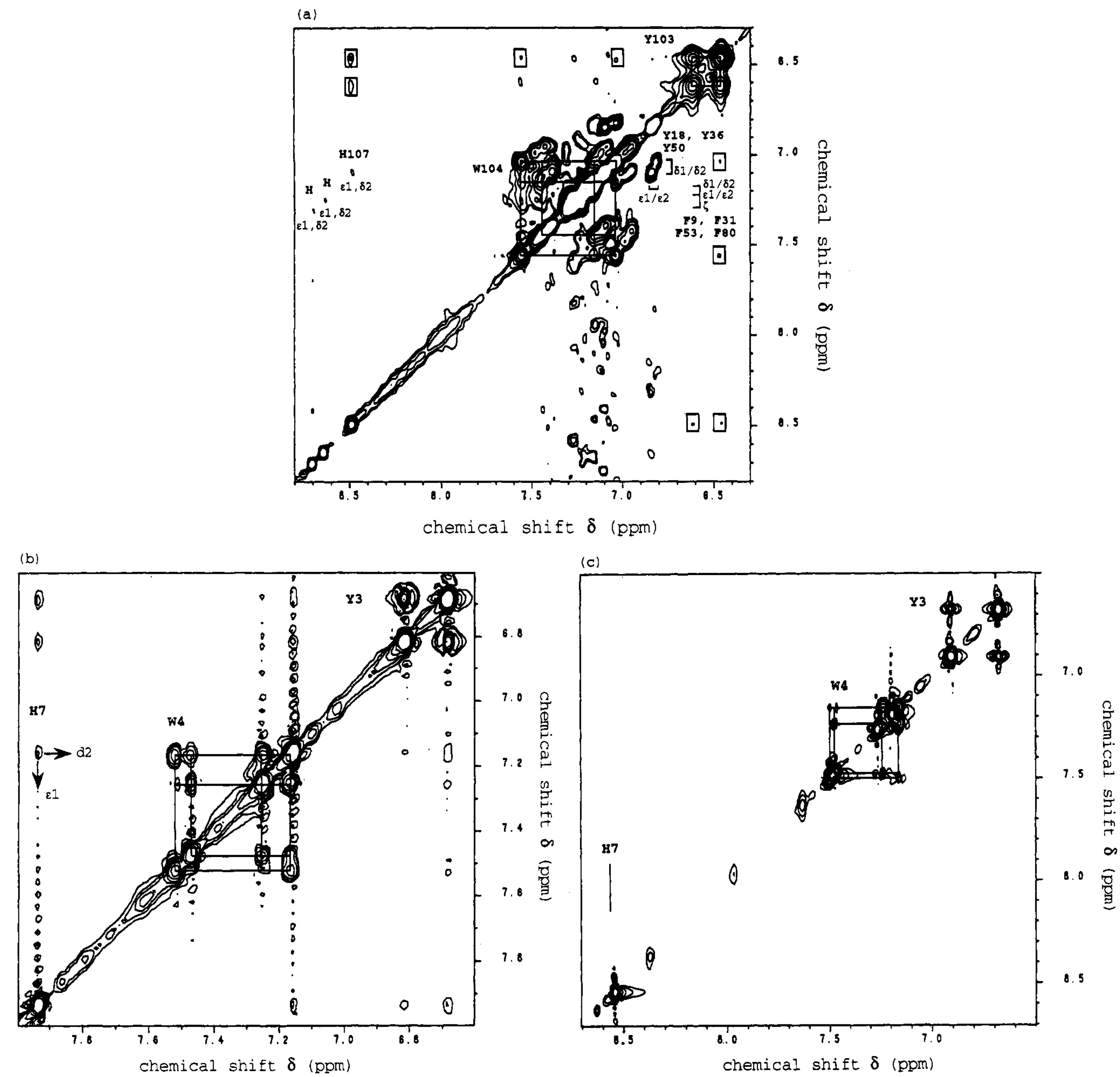

FIGURE 6: Aromatic regions of NOESY spectra of the A-state of BLA and ROESY spectra of the H101–110 peptide in TFE and in water. (a) NOESY spectrum of the A-state of BLA: 3.4 mM protein in D_2O, pH 2.0, 320 K, τ_m = 250 ms, 512 t_1 increments of 2K data points. The data were zero-filled once in ω_2 and three times in ω_1, and multiplied by a Gaussian function (GB = 0.1; LB = −10 Hz) in both dimensions prior to the 2D transform. (b) ROESY spectrum of the H101–110 peptide in TFE: 3.3 mM peptide in 95% CF_3CD_2OD/5% D_2O, pH 2.1, 276 K, τ_m = 150 ms, 620 t_1 increments of 2K data points. The data were zero-filled twice in ω_1 and multiplied by a Gaussian function (GB = 0.1; LB = −4 Hz) in both dimensions prior to the 2D transform. (c) ROESY spectrum of the H101–110 peptide in water: 3.3 mM peptide in D_2O, pH 2.1, 277 K, τ_m = 75 ms, 512 t_1 increments of 2K data points. The data were zero-filled twice in ω_1 and multiplied by a Gaussian function (GB = 0.1; LB = −7 Hz) in both dimensions prior to the 2D transform. Boxed cross-peaks in panel a indicate interresidue NOE effects symmetrical with respect to the diagonal. In addition to NOE cross-peaks, artifact peaks due to a t_1 ridge occur in the region 7.4–6.8 ppm. Note that the intensities of the H107–Y103 NOEs are unsymmetrical with respect to the diagonal. Similar asymmetry is observed in 1D NOE difference spectra, where the NOE on H107 obtained by saturating Y103 is much greater than the NOE on Y103 when H107 is saturated. This asymmetry is likely to reflect different relaxation parameters for Y103 and H107 ring protons. Only negative contour levels are plotted in NOESY spectra; both positive and negative contours are plotted in ROESY spectra.

et al., 1992). If the aromatic residue interactions of the native state were preserved in the A-state of BLA, it should be possible to observe a number of corresponding key NOEs. In the native state of BLA, NOEs are observed between the ε1CH proton of H32 and both the ε1/ε2CH and δ1/δ2CH protons of F31, and between the former and the ζ2CH and η2CH protons of W118. The assignment of the ε1CH proton of H107 in the A-state (see below) requires that the ε1CH proton of H32 resonates at either 8.69 or 8.62 ppm. Two weak cross-peaks in Figure 6a can be attributed to intraresidue NOEs between the ε1CH and the δ2CH protons ring protons of these histidines since the corresponding $^4J_{HH}$ coupling cross-peaks are observed in the same positions in the COSY spectrum of the A-state (see Figure 2b). Neither of the histidine ε1CH protons of H32 or H68, however, exhibits NOE connectivities to aromatic protons from other residues. Similarly, NOEs between δ1/δ2CH and ε1/ε2CH protons of F31 and the ε1/ε2CH protons of Y36 are predicted for cluster I and are observed in NOESY spectra of the native state. The ε1/ε2CH proton resonances of Y18, Y36, and Y50 in the A-state occur between 6.82 and

Table III: Chemical Shift Values (ppm) and Assignments of Resonances from the 101–110 Region of BLA in the A-State[a]

residue	resonance
I 101[b]	γCH_3 0.86; δCH_3 0.77
D 102	
Y 103	αCH 4.21; βCH_2 3.05, 2.81; δ1/δ2CH 6.61; ϵ1/ϵ2CH 6.45
W 104	αCH 4.44; βCH_2 3.42, 3.30; δ1CH 7.18; ϵ1NH 9.99; ζ2CH 7.44; ϵ3CH 7.56; η2CH 7.16; ζ3CH 7.03
L 105[b]	αCH 4.27; γCH 1.76; δCH_3 0.89
A 106	αCH 4.18; βCH_3 1.30
H 107	δ2CH 7.19; ϵ1CH 8.48

[a] Assignments for the ring protons of Y103, W104, and H107 were obtained as described in the text. The assignments of the αCH and βCH resonances of Y103 and W104 are based on standard methods: identification of characteristic strong NOEs between ring protons and αCH and βCH resonances (Wüthrich, 1986). The spin system of A106 (one of three alanine residues in BLA) was readily identified in COSY spectra. The sequence-specific assignment of A106 is based on the observation of similar strong NOEs between this residue and H107 and weak NOEs between A106 and Y103 in both the H101–110 peptide and intact BLA. [b] It has not yet been possible to identify the spin systems of I101 and L105 in COSY spectra of the A-state of BLA. Assignments are based on the presence of NOE cross-peaks in similar positions in ROESY spectra of the H101–110 peptide. Because of this, these assignments should be considered tentative at this time.

6.85 ppm; this group of resonances is relatively well separated from resonances of other types of aromatic protons. The fact that the tyrosine ϵ1/ϵ2CH protons exhibit only intraresidue NOE effects, in the aromatic region of the spectrum, precludes a native-like interaction between Y36 and F31 in the A-state of BLA. Cluster II predicts NOEs between Y103 and W60 and between Y103 and W104. Particularly strong NOEs are observed between the ring protons of Y103 and the Nϵ1 protons of W60 and of W104 in the native state. The tryptophan Nϵ1 protons exhibit only intraresidue NOE effects in the A-state of BLA (Figure 3b). In Figure 6a NOE effects can be discerned between Y103 and only one tryptophan spin system. The presence of similar NOEs in the ROESY spectrum of the H101–110 peptide (Figure 6b) indicates that this tryptophan is W104.

The two clusters of aromatic residues give rise to 26 identifiable NOEs in the native state of BLA. Of these possible connectivities, 17 were certainly not present in NOESY spectra of the A-state (a further eight possible connectivities could not be established because of resonance overlap in the A-state). Only one NOE, between the neighboring residues Y103 and W104, occurs in both the native and A-state of BLA. The absence of NOEs characteristic of cluster I and cluster II strongly suggests that in the A-state these structures are either disordered or different from those in the native state. Of particular significance is the absence of long-range NOEs (e.g., between H32 and W118 in cluster I and between Y103 and W60 in cluster II) since the observation of these would put strong constraints on the overall fold of the chain.

Structure in the 101–110 Region. Resonances from three aromatic residues (a tyrosine, a histidine, and a tryptophan) show particularly large chemical shift deviations from random coil values in the A-state (Figure 1, Table III). Sequence-specific assignments for these well-resolved resonances were obtained by using magnetization transfer spectroscopy (Forsén & Hoffman, 1963; Dobson & Evans, 1984) to correlate resonances from the native and thermally denatured states of BLA: the tyrosine is Y103, the tryptophan is W104, and the histidine is H107 (Alexandrescu et al., 1991). Resonance assignments for Y103, W104, and H107 have been obtained in a similar fashion for the A-states of GPLA (Baum et al., 1989) and HLA (A.T.A. and C.M.D., unpublished results). These assignments were subsequently confirmed by NOESY spectroscopy (see below) of the A-state and supported by studies of a peptide model corresponding to residues 101–110 of HLA (H101–110). Similar chemical shift perturbations are observed for Y103, W104, and H107 in the H101–110 peptide in TFE (Figure 1) suggesting that these deviations are due primarily to a structure localized to a small segment of the chain. For example, the ϵ1CH but not the δ2CH resonance of H107 exhibits a large upfield shift in NMR spectra of the H101–110 peptide in TFE. In the A-states of intact BLA, HLA, and GPLA this resonance shifts to the same position at TFE concentrations greater than 50% (v/v) TFE. The titration behavior is specific to H107, and the positions of the remaining histidines (e.g., H32 and H68 in BLA) do not show perturbations in excess of 0.1 ppm at 50% TFE (Figure 1), indicating that a nonspecific change in the electrostatic properties of histidine residues in the presence of TFE can be ruled out as an explanation for the titration behavior. For the H101–110 peptide in water, the chemical shift values of Y103, W104, and H107 are closer to random coil values although the four resonances indicated in Figure 1 still exhibit deviations from random coil values in excess of 0.15 ppm (compare to the Y18, Y36, and Y50 resonances centered near 7.10 and 6.85 ppm in Figures 1 and 2b).

In the A-states of all three proteins, substantial line broadening was observed for Y103 indicative of conformational averaging of this residue resulting either from mobility within the 101–110 region of the sequence or from motion of the 101–110 segment with respect to the rest of the chain. Line broadening becomes vanishingly small at very high temperatures (above 60 °C), while significant chemical shift deviations from random coil values for Y103 persist. Similarly, in the presence of 50% (v/v) TFE, line broadening is reduced throughout the spectrum of the A-state of BLA while the chemical shift deviations for Y103 remain significant (Figure 1). Decreased affinity for the hydrophobic probe ANS (1,8-anilinonaphthalene sulfonate) as well as increased photo-CIDNP effects for aromatic residues suggest that TFE disrupts the residual tertiary structure of the A-state (A.T.A. and C.M.D., unpublished results). Finally, comparable line broadening is not observed for Y103 in the H101–110 peptide. Taken together, these observations also suggest that the chemical shift deviations for the Y103 resonances can be attributed mainly to interactions within the 101–110 region.

Figures 6 and 7 compare regions from the NOESY spectrum of the A-state of BLA and the corresponding regions of ROESY spectra of the H101–110 peptide in TFE and in water. Direct comparison of these spectra is complicated by the different correlation times of the two molecules and by the different conditions used to obtain cross-relaxation data. It was, for example, necessary to resort to ROESY spectroscopy for the H101–110 peptide since $\omega_0\tau_c \leq 1$. Furthermore, it was valuable to use low temperatures in order to achieve maximal stabilization of structure in the H101–110 peptide, while good quality data for the A-state of BLA could be obtained only at temperatures above 315 K. ROESY spectra of the H101–110 peptide in water (Figure 7c) contain several interresidue NOE cross-peaks, indicating that structured conformations of the peptide are significantly populated. The NOE effects observed for the H101–110 peptide in water are largely a subset of the NOE effects observed for the peptide in 95% TFE (Figure 7b). The larger number of NOEs observed for the H101–110 peptide in 95% TFE is likely to indicate that nonrandom conformations are more populated

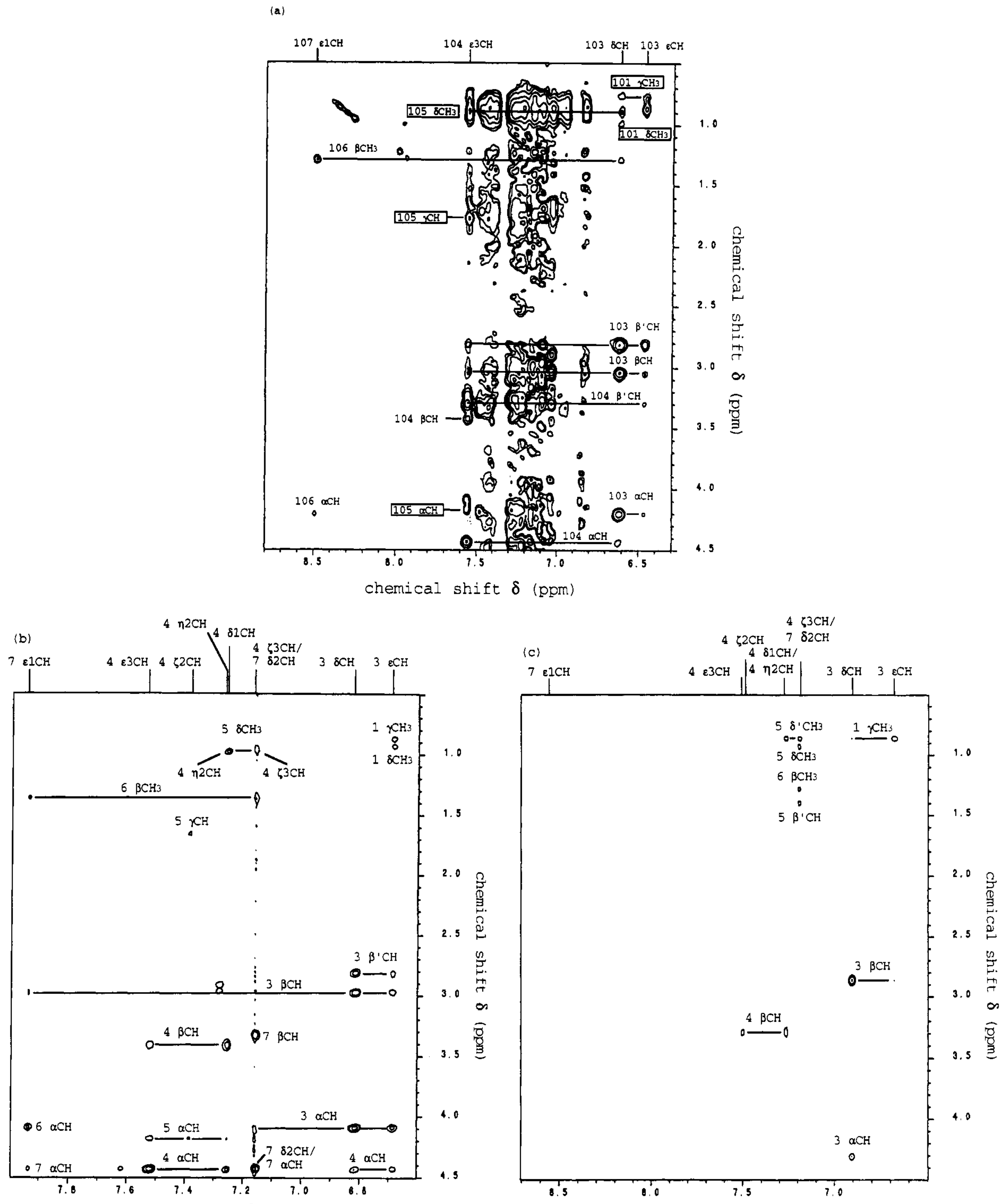

FIGURE 7: Aromatic to aliphatic region of the NOESY spectra of the A-state of BLA and ROESY spectra of the H101–110 peptide in TFE and in water. (a) NOESY spectrum of the A-state of BLA. (b) ROESY spectrum of the H101–110 peptide in 95% CF_3CD_2OD. (c) ROESY spectrum of the H101–110 peptide in water. The conditions are the same as for Figure 6. Spectra are plotted at the same contour levels (intensity) as in Figure 6. NOE assignments indicated by boxes in the NOESY spectrum of BLA were made by comparison to the ROESY spectrum of the H101–110 peptide and should be considered tentative.

for this peptide in TFE than in aqueous solution, consistent with the greater chemical shift dispersion for the peptide in TFE. Because of the limited resolution in the NMR spectrum of the A-state of intact BLA, only the H107 ε1CH proton resonance, the δ1/δ2CH and ε1/ε2CH proton resonances of Y103, and the ε2CH proton resonance of W104 are amenable to comparison with the peptide. Most of the NOE effects observed for these four resonances in the H101–110 peptide in 95% TFE (Figures 6b and 7b) can also be identified in the A-state of intact BLA (Figures 6a and 7a). The A-state of BLA, however, manifests a number of NOE effects not observed in the H101–110 peptide. For example, the W104 ε3CH proton resonance in the A-state exhibits a larger number of NOEs than the corresponding proton in the H101–110 peptide. Since we could only detect three of the four tryptophan Cε3 resonances in 2D NMR spectra of the A-state (Figure 2b), we cannot rule out the possibility that this peak is not completely resolved. Alternatively, the increased number of NOEs for this proton could reflect a different environment for W104 in the peptide and in intact BLA. The fact that some of the subset of NOEs observed in the H101–110 peptide in TFE can also be identified in the A-state of BLA (e.g., Y103 ring protons to H107 ϵ1CH, Y103 ring protons to W104 αCH, Y103 ring protons to I101 γCH and δCH3, H107 ϵ1CH to A106 αCH, H107 ϵ1CH to A106 βCH_3) suggests, however, that the overall fold of the 101–110 sequence is similar in the two forms and that the interactions between Y103, W104, and H107 are to a large extent unaffected by the conformation of the rest of the chain. The small differences in NOE patterns indicate that subtle differences may exist between the peptide and the corresponding region in the A-state. Nevertheless, the H101–110 peptide in TFE appears to be a reasonable model for at least some of the interactions in the 101–110 region of the A-state of BLA.

The NOEs between Y103 and H107 (Figure 6a,b) are of particular significance. In the native-state X-ray structure of HLA (Acharya et al., 1991), the closest distance between the ϵ1CH proton of H107 and the ϵ1/ϵ2CH protons of Y103 is 13.6 Å; that between the ϵ1CH proton of H107 and the δ1/δ2CH protons of Y103 is 11.3 Å. Harata and Muraki (1992) recently isolated a second crystal form of HLA at low-pH (pH 4.2). The X-ray structures of the two forms are virtually identical (e.g., the Ca^{2+}-binding site is conserved in both), except for residues 105–110, which exhibit large differences between the two forms. In the low-pH structure, the closest distance between the ϵ1CH proton of H107 and the ϵ1/ϵ2CH protons of Y103 is 11.4 Å; that between the ϵ1CH proton of H107 and the δ1/δ2CH protons of Y103 is 9.8 Å. Furthermore, NOEs between Y103 and H107 were not detected in NOESY spectra of native BLA recorded at either pH 10.5 (Alexandrescu et al., 1992) or pH 4.3. In addition to the NOEs between the ring protons of Y103 and H107, a number of other NOEs identified in both the H101–110 peptide in 95% TFE and in the A-state of BLA are inconsistent with both X-ray structures (e.g., H107 ϵ1CH–A106 αCH, H107 ϵ1CH–A106βCH_3, Y103ϵ1/ϵ2CH–I101γCH_3, W104ϵ3CH–W104αCH).

pH Dependence of Structure in the 101–110 Region. Figure 8 shows 1D NMR and NOE difference spectra of reduced carboxymethylated BLA (8CM-BLA) as a function of pH. At a temperature of 303 K, 8CM-BLA is denatured in the pH range 2.0–11.0. At low pH, NOEs are observed between Y103 and H107 indicating that the structure in the 101–110 region of the A-state does not depend on the integrity of the four native disulfide linkages, consistent with the disulfide rearrangement experiments of Ewbank and Creighton (1991). Strong support for the proximity of H107 and Y103 comes from the observation that the ionization of H107 induces second-order chemical shift changes in the resonances of the Y103 ring protons (Figure 9). Of the four tyrosine residues in BLA, only Y103 exhibits a titration curve with two pK values: one of 7.1 corresponding to the pK_a of H107, the second of 10.9 corresponding to the pK_a of Y103. The unresolved group of resonances due to Y18, Y36, and Y50 titrates with a single pK_a of approximately 10.9.

The intensity of the NOE between the ϵ1CH proton of H107 and the ϵ1/ϵ2CH protons of Y103 is strongly pH dependent (Figure 8). The NOE becomes undetectable, and the ring proton resonances of Y103 shift closer to random coil values, at pH values above the pK_a of H107 (pH 7.1). While the ionization state of H107 affects the chemical shifts of Y103 resonances (Figure 9), the resonances of H107 are not affected by the ionization state of Y103, indicating that the residues no longer interact at high pH. Furthermore, in contrast to spectra obtained at low pH, ROESY spectra of

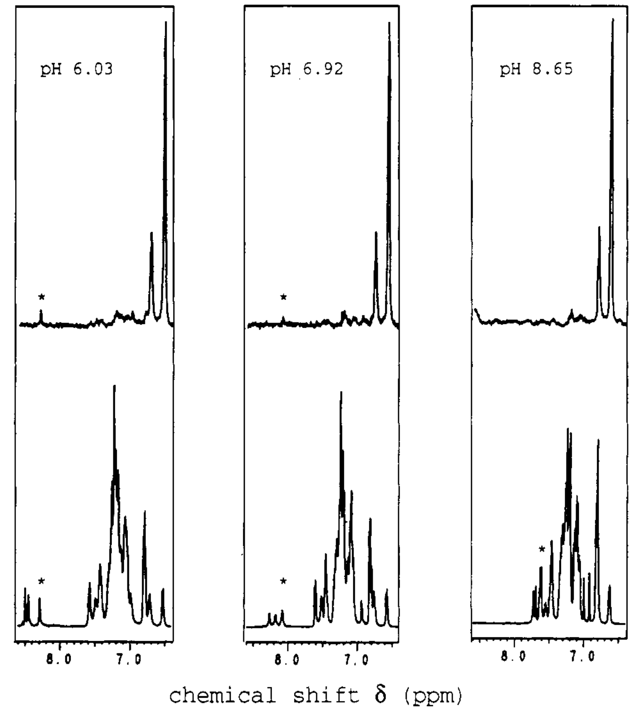

FIGURE 8: pH dependence of an NOE between Y103 and H107 in 8CM-BLA. The bottom traces show 1D NMR spectra of the aromatic region of 8CM-BLA as a function of pH, at a temperature of 298 K. The top traces show 1D NOE difference spectra obtained at the corresponding pH values. The NOE difference spectra were generated by subtracting the FIDs of experiments in which the ϵ1/ϵ2CH resonance of Y103 was saturated from FIDs in which the decoupler was placed in a region of the spectrum which contained no resonances (control spectra). All spectra were normalized to the Y103 ϵ1/ϵ2CH resonance to allow direct comparison of NOE effects. The NOE difference spectra are magnified by a factor of 8 compared to the normal spectra. The asterisk denotes the position of the ϵ1CH proton resonance of H107 in each spectrum. The artifact at 8.60 ppm in the NOE difference spectrum at pH 8.65 is due to off-peak saturation in the control spectrum of this experiment.

H101–110 in 95% TFE at pH 8.3 show only a few interresidue NOEs and no nonsequential interresidue NOEs. Taken together, these observations indicate that the nonnative interaction between Y103 and H107 occurs only when H107 is fully protonated and that this interaction has a similar pH dependence in the H101–110 peptide in TFE. It is worth noting that a pK_a value of 7.1 for H107 in 8CM-BLA implies that the nonnative structure in the 101–110 region of denatured BLA may be substantially populated at physiological pH.

DISCUSSION

The NMR spectrum of the A-state of α-lactalbumin manifests greater chemical shift dispersion than expected for a random coil conformation or indeed found for many other denatured proteins. That the chemical shift dispersion in the A-state is much smaller than in the native state, however, strongly indicates that most of the long-range tertiary structure of the A-state is more disordered and nonspecific than that of the native state (Dolgikh et al., 1985; Baum et al., 1989); the disorder in the tertiary structure of the A-state is illustrated in particular by the absence of NOEs corresponding to the two major aromatic clusters which are present in the native state. The NMR data are consistent with results from a variety of techniques (Dolgikh et al., 1981; Ewbank & Creighton, 1991; Semisotnov et al., 1991) which also indicate that the A-state has a disordered tertiary structure. This leads to an

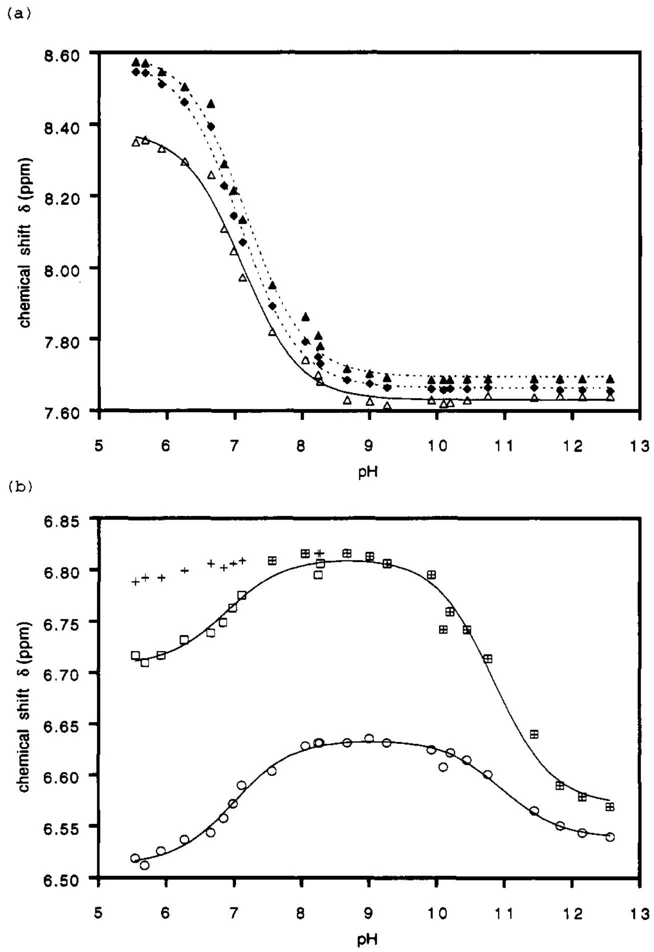

FIGURE 9: pH dependence of tyrosine and histidine resonances in 8CM-BLA. The data were collected on a 3 mM sample of 8CM-BLA at a temperature of 303 K. (a) Histidine Cε1 protons: (△) H107; (▲, ◆) H32, H68. (b) Tyrosine ring protons: (○), ε1/ε2CH protons of Y103; (□) δ1/δ2CH protons of Y103; (+) unresolved ε1/ε2CH resonances of Y18, Y36, and Y50.

apparent paradox since measurements of the hydrodynamic radii of the protein (Dolgikh et al., 1985; Damaschun et al., 1986; Gast et al., 1986) indicate that the A-state is nearly as compact as the native state. Clearly, there must be some force which drives the A-state into a compact form. Theoretical work suggests that a major force contributing to the compactness of denatured states could be the hydrophobic effect (Stitger et al., 1991; Alonso et al., 1991). This is supported experimentally by calorimetric analysis of guanidine hydrochloride denatured BLA (Xie et al., 1991). The upfield chemical shift deviations of side-chain resonances in this work as well as the observation of NOE effects between the side chains of aromatic and aliphatic residues are also consistent with side-chain interactions in the A-state, at least at the level of hydrophobic clustering.

The direction of the chemical shift deviations of αCH proton resonances from random coil values and the observation of protection of amide protons from solvent exchange are consistent with α-helical structure in the A-state detected by other methods (Dolgikh et al., 1985; Kuwajima, 1977, 1989). In the case of GPLA, it has been possible to attribute most of the protected protons in the A-state to regions of the sequence which make up the B and C helices in the native state (Chyan et al., 1992). For BLA we have been able to confirm that at least the C-helix is protected in the A-state although resonances of the B-helix have not yet been assigned in the native state. The presence of stretches of NH–NH NOEs is highly suggestive of helical structure in the A-state of BLA, although the direct identification of helical structure in NOESY spectra of the A-state has not yet proved possible.

In separate work (D. Raleigh, K. Bolin, and C.M.D., unpublished results), peptide fragments corresponding to the helical elements of native α-lactalbumin have been shown to be largely unstructured in water but helical at concentrations in excess of 20% (v/v) TFE. This suggests that the shielding from water provided by the structure of the A-state can, in some respects at least, be modeled by TFE. At high concentrations of TFE, the peptide corresponding to residues 101–110 appears to adopt a conformation related to the corresponding region of the sequence in the A-state of BLA, suggesting that the structure in the 101–110 region of the A-state depends not only on factors local to this region of the sequence but also on the overall environment in which it is formed. The H101–110 peptide in 95% TFE (pH 2.0) shows little or no helical structure by NMR or CD (A.T.A. and C.M. D., manuscript in preparation); thus TFE does not appear to stabilize helical structure nonspecifically, consistent with results obtained in other polypeptide systems (Nelson & Kallenbach, 1989; Lehrman et al., 1990; Segawa et al., 1991). The presence of NOE effects not seen in the native state indicates that structure in the 101–110 region of BLA in the A-state differs from that of the corresponding region in the completely folded protein. In the native state of α-lactalbumin, structure in the 101–110 region is constrained by the tertiary fold of the protein. In particular, the hydrophobic box which includes Y103 and W104 is present in both the high- and low-pH X-ray structures (Harata & Muraki, 1992). Our NMR data indicate that this structure is not conserved in the A-state. The structure observed in the A-state of BLA at pH 2.0 may represent an inherent conformational preference of the 101–110 region which is untenable in the context of the fully folded protein.

The model of the A-state that emerges from this work is of a conformation with a nonspecific tertiary structure which is nevertheless sufficient to lead to a compact conformation and to stabilize elements of localized secondary structure. In the case of the B- and C-helices, the residual structure has features in common with the corresponding regions of the native state. In the case of the 101–110 region of the sequence, however, the structure in the A-state differs from that of the native state. The finding that at least some of the structural characteristics of the A-state can be generated by the addition of TFE to aqueous solutions of peptide fragments suggests that exclusion of water, in the A-state presumably generated by a hydrophobic collapse of the structure, may be crucial in defining the characteristics of the molten globule state.

Further study is needed to obtain a more detailed structural characterization of the A-state. The limited chemical shift dispersion and line broadening in NMR spectra of the A-state severely complicates assignment and analysis. Nevertheless, chemical shift dispersion exists for the A-state, and nonsequential NOE effects can be detected. This implies both that nonrandom structure exists in the A-state and that this structure can be studied by NMR. In many aspects, the problems associated with ^{1}H-NMR studies of the A-state are similar to those encountered in ^{1}H-NMR studies of large proteins. The use of multidimensional NMR in conjunction with isotopic labeling of proteins with ^{13}C and ^{15}N has greatly simplified the interpretation of NMR data of large proteins and is likely to prove extremely useful for NMR studies of denatured proteins (Neri et al., 1992).

ACKNOWLEDGMENT

We thank Anthony Willis for native capillary electrophoresis of α-lactalbumin samples and for amino acid analysis of the H101–110 peptide. Yuet-Lan Ng provided expert technical assistance. We also thank Dr. Kazuaki Harata for providing X-ray coordinates of the low- and high-pH forms of human α-lactalbumin.

REFERENCES

Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M., & Phillips, D. C. (1989) *J. Mol. Biol. 208*, 99–127.

Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., & Fenna, R. E. (1991) *J. Mol. Biol. 221*, 571–581.

Alexandrescu, A. T., Ng, Y.-L., Evans, P. A., Baum, J., & Dobson, C. M. (1991) *J. Cell. Biochem. (Suppl.) 15G*, 187.

Alexandrescu, A. T., Broadhurst, R. W., Wormald, C., Chyan, C.-L., Baum, J., & Dobson, C. M. (1992) *Eur. J. Biochem.* (in press).

Alonso, D. O. V., Dill, K. A., & Stigter, D. (1991) *Biopolymers 31*, 1631–1649.

Anil Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207–213.

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 355-360.

Castellino, F. J., & Hill, R. L. (1970) *J. Biol. Chem. 245*, 417–424.

Chaffote, A., Guillou, Y., Delepierre, M., Hinz, H., & Goldberg, M. E. (1991) *Biochemistry 30*, 8067–8074.

Chyan, C.-L., Wormald, C., Dobson, C. M., Evans, P. A., & Baum, J. (1993) *Biochemistry* (submitted).

Damaschun, G., Gernat, Ch., Damaschun, H., Bychkova, V. E., & Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 226–230.

Dobson, C. M. (1991) *Curr. Opin. Struct. Biol. 1*, 22–27.

Dobson, C. M. (1992) *Curr. Opin. Struct. Biol. 2*, 6–12.

Dobson, C. M., & Evans, P. A. (1984) *Biochemistry 23*, 4267–4270.

Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venjaminov, S. Yu., & Ptitsyn, O. B. (1981) *FEBS Lett. 136*, 311–315.

Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bushuev, V. N., Bychkova, V. E., Gilmaskin, R. I., Levedev, Y. O., Semisotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109–121.

Evans, P. A., Topping, K. D., Woolfson, D. N., & Dobson, C. M. (1991) *Proteins: Struct., Funct., Genet. 9*, 248–266.

Ewbank, J. J., & Creighton, T. E. (1991) *Nature (London) 350*, 518–520.

Fejzo, J., Krezel, A. M., Westler, W. M., Maçura, S., & Markley, J. L. (1991) *Biochemistry 30*, 3807–3811.

Forsén, S., & Hoffman, R. A. (1963) *J. Chem. Phys. 39*, 2892–2901.

Gast, K., Ziwer, D., Welfe, H., Bychkova, V. E., & Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 231–236.

Harata, K., & Muraki, M. (1992) *J. Biol. Chem. 267*, 1419–1421.

Harding, M. M., Williams, D. N., & Woolfson, D. N. (1991) *Biochemistry 30*, 3120–3128.

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

Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6965–6972.

Jeng, M., & Englander, S. W. (1991) *J. Mol. Biol. 221*, 1045–1061.

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

Kuwajima, K. (1977) *J. Mol. Biol. 114*, 241–258.

Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87–103.

Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.

Lehrman, S. R., Tuls, J. L., & Lund, M. (1990) *Biochemistry 29*, 5590–5596.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Mulqueen, P. M., & Kronman, M. J. (1982) *Arch. Biochem. Biophys. 215*, 28–39.

Nelson, J. W., & Kallenbach, N. R. (1986) *Proteins: Struct., Funct., Genet. 1*, 211–217.

Neri, D., Wider, G., & Wüthrich, K. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4397–4401.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Segawa, S.-I., Fukuno, T., Fujiwara, K., & Noda, Y. (1991) *Biopolymers 31*, 497–509.

Semisotnov, G. V., Rodionova, N. A., Razgulayev, O. I., Uversky, V. N., Gripas, A. F., & Gilmanshin, R. I. (1991) *Biopolymers 31*, 119–128.

Stitger, D., Alonso, D. O. V., & Dill, K. A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4176–4180.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222*, 311–333.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.

Xie, D., Bhakuni, V., & Freire, E. (1991) *Biochemistry 30*, 10673–10678.