Local Interactions Drive the Formation of Nonnative Structure in the Denatured State of Human α-Lactalbumin: A High Resolution Structural Characterization of a Peptide Model in Aqueous Solution[†,‡]

Stephen J. Demarest,[§] Yuxin Hua,[§] and Daniel P. Raleigh*[,§,||]

Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, and Graduate Program in Biophysics and Graduate Program in Molecular and Cellular Biology, State University of New York at Stony Brook, Stony Brook, New York 11794-3400

Received February 9, 1999; Revised Manuscript Received April 12, 1999

ABSTRACT: There are a small number of peptides derived from proteins that have a propensity to adopt structure in aqueous solution which is similar to the structure they possess in the parent protein. There are far fewer examples of protein fragments which adopt stable nonnative structures in isolation. Understanding how nonnative interactions are involved in protein folding is crucial to our understanding of the topic. Here we show that a small, 11 amino acid peptide corresponding to residues 101−111 of the protein α-lactalbumin is remarkably structured in isolation in aqueous solution. The peptide has been characterized by ^{1}H NMR, and 170 ROE-derived constraints were used to calculate a structure. The calculations yielded a single, high-resolution structure for residues 101−107 that is nonnative in both the backbone and side-chain conformations. In the pH 6.5 crystal structure, residues 101−105 are in an irregular turn-like conformation and residues 106−111 form an α-helix. In the pH 4.2 crystal structure, residues 101−105 form an α-helix, and residues 106−111 form a loopike structure. Both of these structures are significantly different from the conformation adopted by our peptide. The structure in the peptide model is primarily the result of local side-chain interactions that force the backbone to adopt a nonnative 3_{10}/turn-like structure in residues 103−106. The structure in aqueous solution was compared to the structure in 30% trifluoroethanol (TFE), and clear differences were observed. In particular, one of the side-chain interactions, a hydrophobic cluster involving residues 101−105, is different in the two solvents and residues 107−111 are considerably more ordered in 30% TFE. The implications of the nonnative structure for the folding of α-lactalbumin is discussed.

There are many examples of conformational studies of peptide fragments of proteins. These studies provide clues about the propensity to form local structure in the absence of tertiary interactions and, hence, can provide information about the conformational tendencies of the unfolded state. In most cases, small peptides are found to be largely devoid of structure. In some cases, however, a propensity to adopt nativelike structure has been observed and this has led to considerable discussion about the role of local interactions in folding (*1*). There are notably fewer examples of local interactions leading to the formation of significant nonnative structure (*2*). In some cases, nonnative interactions have been shown to be slow folding (*3*−*6*). This raises the question of what impact local and nonlocal nonnative interactions have upon the correct folding of the polypeptide chain to the native state. Structural characterization of nonnative interactions that exist in denatured or partially folded states of proteins is a necessary first step in deconstructing their involvement in the folding process.

Here we have used two-dimensional ^{1}H NMR[1] to analyze a peptide derived from residues 101−111 of the protein α-lactalbumin (αLA) in order to probe for local structural preferences that may play a role in stabilizing partially folded states. A high-resolution analysis of the nonnative structure found in this peptide is described. This is a particularly interesting region of αLA. It adopts very different conformations in the crystal structures of the high pH (6.5) and low pH (4.2) forms (*7*, *8*). In the pH 6.5 crystal structure (PDB accession code 1HFX), residues 101−105 are in an irregular

[†] This work was supported by NIH Grant GM54233 to D.P.R. D.P.R. is a Pew Scholar in the Biomedical Sciences. S.D. was supported in part by a GAANN Fellowship from the Department of Education. The NMR facility at SUNY Stony Brook is supported by a grant from the NSF CHE9413510.

[‡] Atomic coordinates for the 40 NMR structures of αlac:101-111 in aqueous solution, together with listings of the experimental restraints used in this study, have been deposited in the Protein Data Bank at Rutgers (access code 1CB3).

* To whom correspondence should be addressed. E-mail: draleigh@ccmail.sunysb.edu. Phone: 516-632-9547. Fax: 516-632-7960.

[§] Department of Chemistry, SUNY at Stony Brook.

[||] Graduate Program in Biophysics and Graduate Program in Molecular and Cellular Biology, SUNY at Stony Brook.

[1] Abbreviations: CD, circular dichroism; Da, daltons; DQF-COSY, double quantum filtered correlated spectroscopy; E.COSY, exclusive correlated spectroscopy; Fmoc, 9-flourenylmethyloxycarbonyl; HPLC, high-pressure liquid chromatography; NMR, nuclear magnetic resonance; PAL, 5-(4′-Fmoc-aminomethyl-3′,5′-dimethoxyphenoxy)valeric acid; RMSD, root-mean-squared deviation; ROESY, rotating frame nuclear Overhauser effect spectroscopy; TBTU, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate; TFA, trifluoroacetic acid; TFE, trifluoroethanol; TOCSY, total correlated spectroscopy; TSP, 3-(trimethylsilyl) propionate.

10.1021/bi990320z CCC: $18.00
Published on Web 05/14/1999

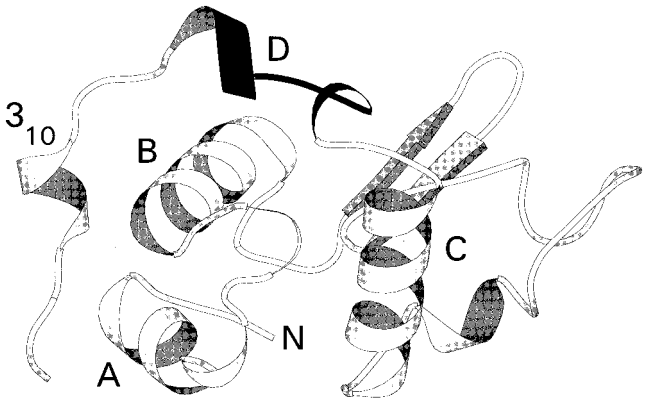

FIGURE 1: MOLSCRIPT ribbon diagram of the X-ray crystal structure of human αLA (32). The α-domain is in white and the *β*-domain is in gray. Residues 101−111 are shaded black.

turn-like conformation and residues 106−111 form an α-helix. In the crystal structure determined at pH 4.2, residues 101−105 form an α-helix, and residues 106−111 form a looplike structure. This region of αLA also appears to play an important role in stabilizing the molten globule state. Proline scanning mutagenesis studies suggest that the D-helix (residues 106−111) is formed in the molten globule state (*9*), and protein dissection studies have demonstrated that residues 101−111 are required for molten globule formation (S.J.D., J. Boice, R. Fairman, and D.P.R., unpublished results).

αLA consists of two subdomains, a primarily helical domain (α-domain) consisting of residues 1−39 and 81−123 and a sheet/coil domain (*β*-domain) encompassing the remainder of the molecule (Figure 1) (*7*). In the native state of αLA, the α-domain contains four major α-helices, denoted A−D, and a C-terminal 3_{10} helix. We have previously obtained qualitative evidence for nonnative structure in isolated peptide fragments of αLA containing the D helix and the C-terminal 3_{10} helix (*10*). One of the nonnative interactions characterized in the D-helix region of the isolated peptides, a Tyr−His interaction between residues 103 and 107, has been shown to exist in the acid-denatured molten globule state of the full-length protein (*11*). In the homologous protein, hen lysozyme, there are two Trp residues corresponding to positions 104 and 107 of the αLA sequence. In lysozyme, these residues have been shown to slow folding by stabilizing an intermediate through nonnative interactions (*3*).

In our previous study, we used CD and NMR to investigate the conformational tendencies of a peptide corresponding to residues 101−111 of αLA (designated αlac:101−111) (*10*):

^{101}I D Y W L A H K A ^{110}L A

In the present work, we have collected additional ^{1}H NMR data and performed detailed calculations of the structure of the peptide in H_2O and in 30% (v/v) TFE/H_2O. The peptide is remarkably structured in the two solvents, and the observed structures are notably different from the structure calculated previously for a similar peptide in 95% TFE (*12*). We compare the structures of our peptide in the two solvents and discuss the insights our structures provide for understanding the role of this region in the folding of αLA.

MATERIALS AND METHODS

Peptide Synthesis and Purification. The native Cys residue at position 111 of αlac:101−111 was replaced by Ala. Peptide synthesis reagents and solvents were purchased from PerSeptive Biosystems Inc. and Fisher Scientific, respectively. αLac:101−111 was synthesized on a Rainin model PS3 automated peptide synthesizer on a 0.1 mmol scale. Fmoc-L-amino acids were coupled via TBTU activation. The use of a polystyrene support with a peptide amide linker (PAL resin) provided an amidated C-terminus. The N-terminus was acetylated upon completion of the synthesis. Cleavage from the resin was accomplished using a solution of 90% TFA, 3.3% ethanedithiol, 3.3% anisole, and 3.3% thioanisole. The peptide was purified by reversed-phase HPLC using a Vydac C18 column with H_2O/CH_3CN gradients containing 0.1% TFA. The peptide was characterized by MALDI mass spectrometry; expected molecular mass, 1341.5 Da, observed molecular mass, 1341.0 Da.

CD. Measurements were performed on an AVIV model 62A circular dichroism spectrometer. Wavelength scans were performed with an averaging time of 3 s and a minimum of three repeats. Peptide concentrations were 100 μM. A 2 mM phosphate, 2 mM borate, 2 mM citrate, and 10 mM NaCl buffer at pH 2.8 was used for all CD experiments. TFE was purchased from Aldrich. Peptide concentrations were determined by absorbance measurements using extinction coefficients calculated via the method of Pace and co-workers (*13*).

^{1}H NMR. The samples used for experiments in H_2O were 3 mM αlac:101−111 in 90% H_2O and 10% D_2O (v/v) at pH 2.8 (unbuffered). NMR samples containing TFE were prepared by first dissolving αlac:101−111 in H_2O. The pH was adjusted to 2.8 and TFE-d_3 was added to a final concentration of 30% (v/v). The final peptide concentration was 3 mM. All deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. NMR experiments were carried out on Varian Unity INOVA 500 and 600 MHz spectrometers. Presaturation was used for solvent suppression. TOCSY and ROESY experiments were performed with either 256 or 512 t_1 increments and 2048 data points in t_2. TOCSY experiments utilized a 75 ms mixing time, and ROESY experiments used a 300 ms mixing time. DQF-COSY and E.COSY experiments were performed with 512 t_1 increments and 8K and 4K data points in t_2, respectively. All experiments used the TPPI method of data collection (*14*) except the E.COSY experiments, which used the method of States and co-workers (*15*). All data sets were processed using Felix 95.0 (Biosym). Sinebell window functions shifted 70−90° were applied to all 2D data sets. The data sets were zero-filled once in the first dimension and twice in the second dimension before Fourier transformation. Sequence-specific assignments were determined by assigning spin systems from the TOCSY spectra and using sequential alpha to amide ROE peaks (*16*). $^3J_{NH\alpha}$ coupling constants were estimated by measurement of v_a and v_d from DQF-COSY experiments (*17*). $^3J_{\alpha\beta}$ coupling constants were measured as passive couplings from E.COSY experiments (*18*). Intensities of ROESY cross-peaks were corrected for resonance offset effects (*19*). A listing of ^{1}H resonance assignments for αlac:101−111 have been submitted to the BioMagRes Database (Madison, WI).

Distance Restraints and Structure Calculations. The restraints for the calculation of the peptide structure in aqueous solution were generated from experiments performed at 10 °C. All of the restraints used for the structure calculations of the peptide in 30% TFE were derived from experiments performed at 25 °C. ROEs were assigned as strong, medium, weak, and very weak by counting contour levels. Upper boundary limits were set at 3.3 Å for strong, 3.8 Å for medium, 4.5 Å for weak, and 5.0 Å for very weak. Standard psuedoatom distances were used for those methyl and aromatic protons that could not be stereospecifically assigned (*16*). ϕ angles were restricted to $-65 \pm 30°$ for measured $^3J_{NH\alpha}$ values below 6 Hz. χ^1 values were restrained to particular rotamers as determined by the $^3J_{\alpha\beta}$ coupling constants and amide to β and α to β proton ROE intensities (*20*). No hydrogen-bonding restraints were used in any of the calculations.

Structures were calculated using the program X-PLOR version 3.851 (*21*). The standard files, parallhdg.pro and topallhdg.pro, were used to define the force constants and residue topologies, respectively. Random coordinate files were generated using random.inp (*22*, *23*). For both the aqueous solution and 30% TFE data, 200 substructure embedded coordinate files were generated using the dg_sub_embed.inp protocol (*24*, *25*). Distance geometry and simulated annealing was performed on the 200 subembedded structures using the dgsa.inp protocol (*24*−*26*). This protocol performs initial energy minimization using the conjugate gradient method, followed by 1000 steps of restrained molecular dynamics at 2000 K and 1000 steps of simulated annealing to 100 K for a total of 6 ps of dynamics. These structures were then subjected to a final step of minimization. The resulting 200 coordinate files were refined using the protocol refine.inp. Dynamics were begun at 1000 K and consisted of 3000 cooling steps to 100 K for a total of 9 ps of simulated annealing. A final stage of minimization was performed using a repel value of 0.75 to define the VDW potential.

Using a cutoff of 0.2 Å for upper bound ROE violations and 2.5° for dihedral angle bound violations, 40 structures were selected from the 200 refined structures to make up the final acceptable ensemble for αlac:101−111 in aqueous solution. Using the same criteria, 41 refined structures out of 200 passed the violations screen and were chosen to represent the structure of αlac:101−111 in 30% TFE. An average structure was generated from each ensemble using the protocol average.inp. These structures were then subjected to a 4000 step energy minimization using the conjugate gradient method in X-PLOR and are denoted the average/minimized structures. For the assignment of backbone hydrogen bonds, the following criteria were used: the amide proton and carbonyl oxygen must be within 2.4 Å of one another and the deviation from linearity must not exceed 50°.

RESULTS

αLac:101−111 in Aqueous Solution: A Model of Local Interactions in the Denatured State. The CD spectrum of αlac:101−111 at pH 2.8 does not change significantly between 1 and 60 °C. ROESY spectra collected at 5, 10, and 25 °C are all very similar; although, the resolution is best at 10 °C (*10*). At 10 °C, the intesities of the ROE peaks were slightly greater than those at room temperature and the improved spectral resolution allowed for a slightly greater number of ROEs to be assigned compared to other temperatures. Therefore, the data at 10 °C was chosen for further analysis. The similarities in the spectra recorded at 5, 10, and 25 °C lead us to believe that the structure is very similar at these temperatures. αLac:101−111 displays a surprising amount of structure for an 11 residue peptide in aqueous solution. A total of 170 proton ROEs were assigned and used for the structure calculations. Of these ROEs, 136 are between protons separated by at least four bonds. The remaining ROEs were used to differentiate between distances that would give rise to clear differences in ROE intensities (for example, between the α proton of a residue and its two β protons). Of the 170 ROE restraints, 138 involve residues 101−107. The ROE intensities and four backbone dihedral angle constraints were used as restraints in the structure calculations. No hydrogen-bonding restraints were used. An ensemble of 40 structures with minimal restraint violations was generated. The distribution of ROEs and the backbone and heavy atom RMSDs are shown in Figure 2, and statistics for the ensemble are shown in Table 1.

The ensemble displays a very low RMSD for residues 101−107 where the majority of the ROE restraints lie (Figure 3). If only residues 101−107 are superpositioned, the RMSD of the backbone and heavy atoms is 0.28 and 0.54 Å, respectively. These results suggest that residues 101−107 predominately populate one conformation which may be in fast exchange with the random coil state.

The backbone of residues 103−106 adopts a conformation resembling a 3_{10} helix which is terminated after H107. An (i,i+3) hydrogen bond between the carbonyl oxygen of Y103 and amide proton of A106 is present in every member of the ensemble of structures. Another (i,i+3) hydrogen bond between the carbonyl oxygen of W104 and the amide proton of H107 was observed in 16 of the structures. In nine structures, the amide proton of K108 acts as a hydrogen bond donor. In six of these structures, the W104 carbonyl forms a hydrogen bond to the amide protons of both H107 and K108. Amide ^{1}H temperature coefficients well below random coil values (−4.2 and −3.2 ppb/K) were observed for the amide protons of A106 and H107 and are consistent with either hydrogen bond formation or protection from solvent. Residues 103−106 all cluster in the helical region of the Ramachandran plot. The ϕ,ψ angles of D102 fall outside of the normally allowed regions of the Ramachandran plot. In an analysis of protein crystal structures, Moult and Herzberg found that many of the nonglycine residues that deviate from the normally allowed regions of the Ramachandran plot are aspartic acid residues (*27*). Aspartic acid residues are often found as capping residues in helices. However, there is no evidence that D102 functions as a capping residue in this peptide. The conformation of residues 108−111 is not well-defined since there are very few experimental restraints in this region of the peptide.

The observed ROEs provide evidence for the formation of a hydrophobic cluster involving the side chains of I101, Y103, and W104. In all of the structures, the side chains of these residues are clustered together and are in van der Waals contact with one another. There is also an interaction involving the side chains of Y103 and H107. Four medium strength ROEs between the imidazole side chain of H107

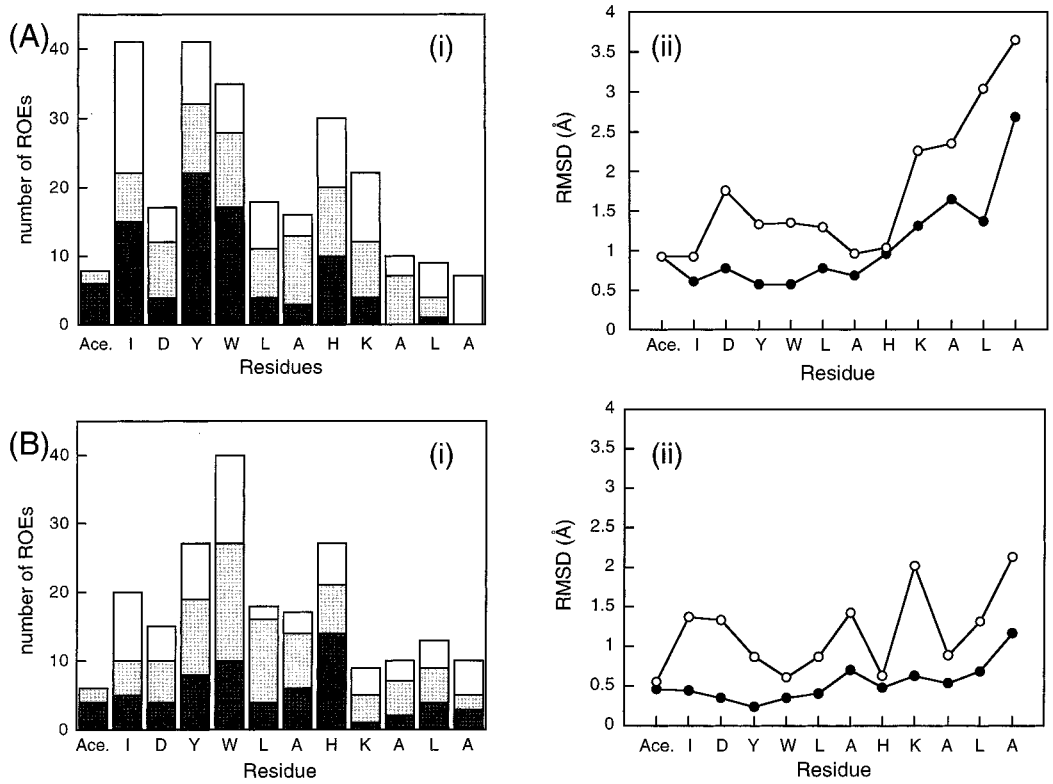

FIGURE 2: Plots of the number of ROEs per residue and the RMSD per residue for (A) the ensemble of structures in aqueous solution and (B) the ensemble of structures in 30% TFE. (i) The ROEs per residue (open, intraresidue; gray, sequential; black, medium range). (ii) The RMSD of each residue (closed circles, backbone; open circles, all heavy atoms).

Table 1: Statistics for the Ensembles of Structures of αlac:101−111 in Aqueous Solution (40 Structures) and in 30% TFE (41 Structures)

	αlac:101−111 in aqueous solution, pH 2.8	αlac:101−111 in 30% TFE, pH 2.8
RMS deviations of the ensemble of structures on the average structure (Å)		
backbone, residues 101−111	1.34 ± 0.32	0.74 ± 0.17
heavy atom, residues 101−111	1.48 ± 0.34	0.98 ± 0.20
backbone, residues 101−107	0.28 ± 0.15	0.32 ± 0.13
heavy atom, residues 101−107	0.54 ± 0.09	0.66 ± 0.08
RMS deviations from experimental restraints		
ROE restraints (Å)	0.026 ± 0.002	0.014 ± 0.004
Dihedral restraints (°)	0.45 ± 0.07	0.87 ± 0.08
RMS deviations from idealized geometry		
bonds (Å)	0.0035 ± 0.0002	0.0023 ± 0.0003
angles (°)	0.59 ± 0.02	0.49 ± 0.03
impropers (°)	0.35 ± 0.02	0.43 ± 0.03
potential energies (kcal mol^{-1})		
bond lengths	2.38 ± 0.21	1.05 ± 0.31
bond angles	18.85 ± 1.39	12.93 ± 1.34
improper angles	2.14 ± 0.26	3.27 ± 0.51
VDW contacts	9.09 ± 1.05	2.67 ± 1.56
dihedral restraints	0.10 ± 0.03	0.35 ± 0.11
ROE restraints	5.76 ± 1.07	1.51 ± 0.85
total	38.17 ± 2.55	21.02 ± 4.19

and the side chain of Y103 are observed. An additional two medium/weak ROEs are observed from the side chain of H107 to the $C_\alpha H$ of Y103. The histidine side chain occupies the g^+ rotamer. This positions the side chain so that it projects down the vertical axis of the helix toward the N-terminus. The conformation of the Y103 side chain is also well-defined, populating the g^- rotamer even though no χ^1 restraints were included in the structure calculations.

None of the residues involved in the hydrophobic cluster have $^3J_{\alpha\beta}$ coupling constants indicative of populating a single rotamer. Averaging with the random coil state of the peptide, which is fast on the NMR time scale, is likely the cause of the averaged coupling constants. The intensities of the ROE cross-peaks which define the position of these side chains will be dominated by the population of molecules which are structured (*2*). In contrast, the coupling constants reflect an average over both the structured and random coil populations. There may also be some variability in the position of the side chains in the hydrophobic cluster that could contribute to averaging of the coupling constants. Some fluctuations in the position of these side chains is reasonable since hydrophobic interactions are nondirectional.

The structure in residues 101−107 is highly nonnative. It differs greatly from the structure of this region of αLA in both the high and low pH crystal forms (*7*, *8*). Superposition of residues 101−106 of the average/minimized structure with

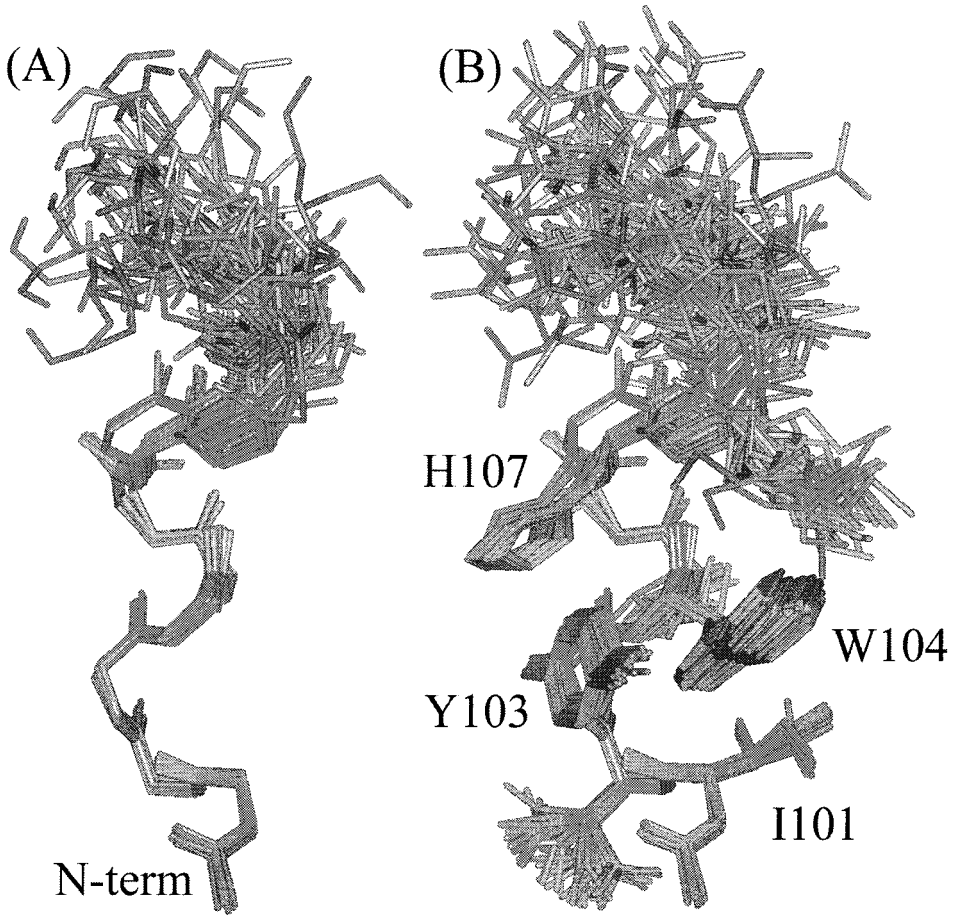

FIGURE 3: (A) Backbone representation and (B) heavy atom representation of the superposition of the 40 structures of the ensemble determined at pH 2.8, 10 °C in aqueous solution.

the same residues of the high pH crystal structure of intact αLA reveals significant differences. The RMSD of the backbone is 2.36 Å, and the RMSD of the heavy atoms is 3.69 Å. A comparison between residues 101−111 from the high pH crystal structure and αlac:101−111 is shown in Figure 4. A similar comparison was not possible for the low pH crystal structure since the coordinates are not publicly available. However, residues 101−105 form an α-helix in the low pH form, and the side chains of Y103 and H107 are over 10 Å apart. This is clearly very different from what is observed in the structure of αlac:101−111 in aqueous solution.

αLac:101−111 in 30% TFE: The Nonnative Structure Differs from That Found in Aqueous Solution. The structures of numerous peptides have been determined in mixed TFE/aqueous solutions (*12*, *28*, *29*). In most cases, peptides are largely unstructured in water and the addition of TFE usually induces helical structure. αlac:101−111 offers an interesting system with which to examine the conformational rearrangements induced in an already structured peptide upon the addition of TFE. Intact αLA has been shown to populate a partially folded state in mixed TFE/H_2O solvents, and thus it is also of interest to examine how TFE modulates the tendency to adopt locally stabilized structure. A similar peptide has previously been studied in 95% TFE; however, it is known that the properties of TFE water mixtures vary dramatically between mixtures which are rich in TFE versus those which are low in TFE (*30*, *31*). A large increase in the far UV CD signal at 222 nm is observed with the addition

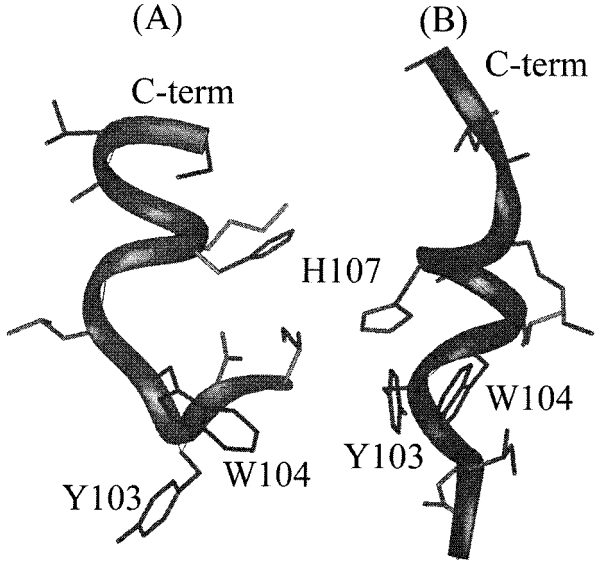

FIGURE 4: (A) Ribbon diagram of residues 101−111 from the high pH crystal structure of human αLA. (B) Ribbon diagram of αlac:101−111 in aqueous solution. The side chains are included to show the large difference between the native structure and the structure of αlac:101−111.

of TFE to an H_2O solution containing αlac:101−111 at pH 2.8, 25 °C (Figure 5). The isodichroic point at 202 nm indicates that each residue is sampling one of two conformations, most probably helix and coil. The intensity at 222 nm reaches its maximum at 30% (v/v) TFE.

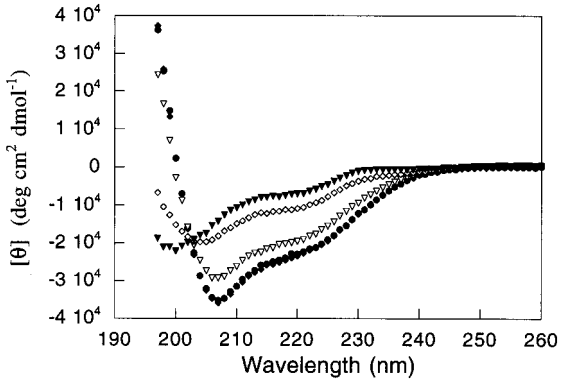

FIGURE 5: TFE titration monitored by CD of αlac101−111 at pH 2.8, 25 °C (filled triangles, 0.0% TFE; open diamonds, 9.9% TFE; open triangles, 19.4% TFE; filled diamonds, 28.6% TFE; filled circles, 40.0% TFE). Note, the 28.6% and the 40% TFE spectra are essentially superimposable.

Several key proton resonances overlap at 10 °C. These resonances are resolved at 25 °C; therefore, the structures were calculated using the data collected at 25 °C. The ROEs which could be assigned at 10 °C were all present in the 25 °C spectra, suggesting that the structures at 10 and 25 °C are very similar. Calculations of the peptide structure in 30% TFE were performed using a total of 138 ROE and 7 dihedral angle restraints. Again, no hydrogen-bonding restraints were used. An ensemble of 41 structures with minimal restraint violations was generated. The overall RMSD is much lower than what is observed for the peptide in aqueous solution: 0.74 Å for all backbone atoms and 0.98 Å for all heavy atoms (Figure 6). However, when only residues 101−107 are superimposed, the RMSDs are comparable (0.32 Å for the backbone atoms and 0.66 Å for the heavy atoms in 30% TFE versus 0.28 Å for the backbone atoms and 0.54 Å for the heavy atoms in H_2O). The distribution of ROEs and plots of the backbone and heavy-atom RMSDs are shown in Figure 2. The statistics for the ensemble of 41 structures are provided in Table 1.

The conformation of the N-terminal portion of the backbone changes very little with the addition of 30% TFE (Figure 7A). If residues 101−106 of the backbone from the average/minimized structure in aqueous solution are superimposed on the backbone from the average/minimized structure in 30% TFE, the RMSD is 0.98 Å. Residues 103−106 still form a 3_{10} helix/turn-like conformation. The Y103 carbonyl group maintains its hydrogen bond with A106 in every member of the ensemble. The carbonyl of W104 is close to the amide protons of both H107 and K108, similar to what is observed in aqueous solution. W104 makes a hydrogen bond with H107 in 19 of the 41 structures and with K108 in 2 of the 41 structures. The greatest difference in the backbone conformation is that the residues C-terminal to H107 are disordered in H_2O, but form an α-helix in 30% TFE. The ϕ,ψ angles of residues 103−111 from the average/minimized structure in 30% TFE all fall in the helical portion of the Ramachandran plot. $i,i+4$ hydrogen bonds are observed between L105 and A109 in 4 of the 41 structures, between A106 and L110 in 3 structures, and between H107 and A111 in 12 structures. All four $i,i+4$ hydrogen bonds

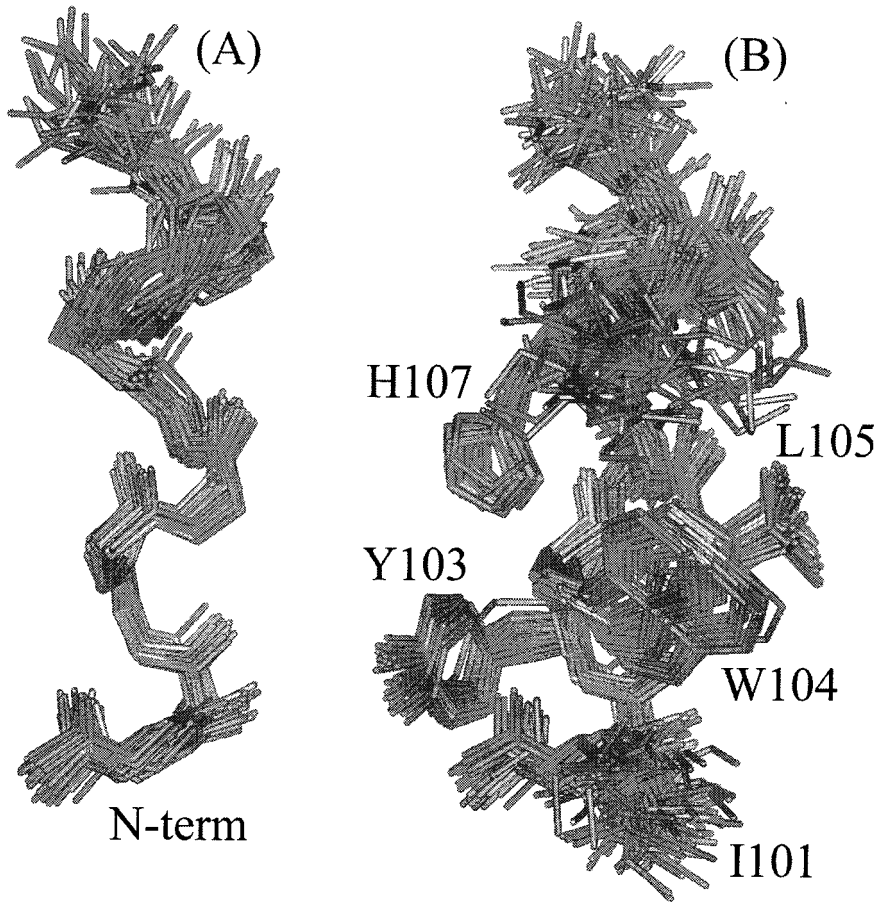

FIGURE 6: (A) Backbone representation and (B) heavy atom representation of the superposition of the 41 structures of the ensemble determined at pH 2.8, 25 °C in 30% TFE.

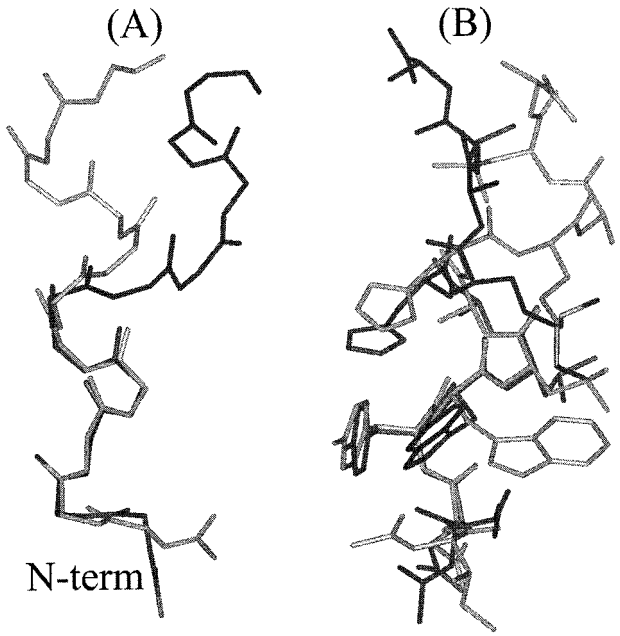

FIGURE 7: Superposition of residues 101−107 of the average/minimized structure at pH 2.8, 10 °C in aqueous solution (black) with residues 101−107 of the average/minimized structure at pH 2.8, 25 °C in 30% TFE (grey). (A) Backbone. (B) Heavy atom.

(W104−K108, L105-A109, A106-L110, and H107-A111) are present in the average/minimized structure. An α-helix is observed from residues 106 to 111 in the high pH crystal structure (*7*). If the average/minimized backbone of residues 106−111 is superimposed upon the backbone of the same residues from the crystal structure, the RMSD is 0.80 Å. In contrast, the backbone RMSD between residues 101−106 of the average/minimized structure and the same residues of the crystal structure is 2.46 Å. Thus, while the backbone conformation of residues at the N-terminus is clearly nonnative, the backbone of residues 106−111 does adopt a nativelike configuration in 30% TFE.

Although the backbone conformation of residues 101−106 in 30% TFE is very similar to that observed in H_2O, the side chain positions are significantly different (Figure 7B). In particular, the hydrophobic cluster loses its integrity. Only one weak ROE is observed between the indole ring of W104 and the phenol side chain of Y103 in 30% TFE compared to five medium strength ROEs in aqueous solution. As a consequence, the conformation of the W104 side chain is no longer constrained by the Y103 side chain. There are also fewer ROEs between Y103 and I101. Additional ROEs are observed between the W104 and L105 side chains in 30% TFE that were not observed in aqueous solution. As a result, the W104 side chain occupies the g^- rotamer in 30% TFE compared to the g^+ rotamer in aqueous solution. This change creates a hydrophobic cluster on one face of the 3_{10}/turn structure which includes the side chains of I101, W104, and L105.

The side chain interaction between Y103 and H107 is maintained in 30% TFE. The same set of six ROEs which were observed in H_2O were observed in 30% TFE and are of the same intensity. In all of the structures, these two side chains adopt the same relative orientation as they do in aqueous solution. It appears that this robust interaction helps to stabilize the 3_{10}/turn-like conformation in residues 103−106.

DISCUSSION

Our calculations have shown that residues 101−107 of αlac:101−111 adopt a well-defined structure in aqueous solution. Although there is doubtless some conformational flexibility, likely including some coiling and uncoiling of the peptide backbone between residues 101−107, a single structure is consistent with all of the ROE restraints. The structure appears to be stabilized by a number of nonnative side chain interactions: the hydrophobic cluster involving residues I101, Y103, and W104 and the interaction between residues Y103 and H107. Since this structure is stable in the absence of tertiary interactions and is the result of local interactions, it will almost certainly be present in the unfolded state at the onset of the folding process at low pH. The structure in our peptide is nonnative and a drastic rearrangement would be required to fold to the low (pH 4.2) or high pH (pH 6.5) crystal structures (*7*, *8*). In the pH 6.5 crystal structure, H107 is buried and makes a hydrogen bond with the carboxyl group of E25 (*7*, *8*). At pH 4.2, H107 is exposed to solvent. The protonated side chain of H107 also appears to be important for the formation of the local structure observed in isolation at low pH and may, thus, preferentially contribute to the stability of the unfolded state at low pH (*10*).

The local structure of αlac:101−111 is particularly interesting because it appears to mimic some of the structure formed in this region of intact αLA in the molten globule state. CD studies have shown that the molten globule state of αLA contains a high degree of secondary structure; however, its dynamic nature and tendency to associate at high concentrations have made detailed structural studies of this state difficult. Dobson and co-workers have used ^1H NOEs to show that the side chains of Y103 and H107 are in close proximity in the acid-denatured molten globule state (*11*). Our study has shown that one of the few well characterized examples of structure in the molten globule state can be induced by local interactions.

Molten globule-like states of proteins generally contain extensive secondary structure and bury a significant number of hydrophobic groups. In intact αLA, residues 101−111 are involved in interactions which are important for the stability of the molten globule state (S.J.D., J. Boice, R. Fairman, and D.P.R., unpublished results). It is possible that the ability of this region to form locally stabilized structures which includes the clustering of hydrophobic side chains may play a role in the formation of the molten globule state. In H_2O, our peptide is essentially unstructured from residues 107 to 111. However, proline-scanning mutagenesis studies have provided indirect evidence that this region adopts helical structure in the molten globule state (*9*). Thus, it is very interesting that in 30% TFE the peptide does form a nativelike helix from residues 106 to 111 while still maintaining the nonnative Y103−H107 side-chain interaction. The hydrophobic cluster including residues I101, Y103, W104, and L105 is also able to rearrange depending upon the environment of the peptide without disrupting the interaction between the side chains of Y103 and H107 that is known to be present in the molten globule state. These observations demonstrate that the nonnative structure observed in the N-terminal portion of our peptide is comparable with the existing structural data for the αLA molten globule

state and shows, but does not prove, that it could be preserved in the molten globule state (*11*).

It is interesting to note that the structure in aqueous solution and the structure in 30% TFE are both very different from the structure of a slightly shorter peptide in 95% TFE (*12*). In 95% TFE, the 3_{10}/turn conformation of residues 103−106 found in our structures is not formed, and there is not as much evidence for the formation of an α-helix in the five C-terminal residues. Instead, a turn-like conformation in residues 105−108 is prevalent in 95% TFE. In the structures determined in both 30% TFE and aqueous solution, the longest range side chain to side chain interactions involved residues which are four apart in primary sequence. Longer range side-chain interactions were reported in 95% TFE, including interactions between residues differing by up to five and six residues in primary sequence. The properties of TFE rich solutions have been shown to be very different from mixed aqueous TFE solutions which contain a low mole fraction of TFE (*30*, *31*). These differences lead to the interesting changes in the peptide structure described here, a result which highlights the potential complications of studies of peptides in mixed solvents.

In summary, the comparison of the structure of our peptide in aqueous solution with the structure in 30% TFE shows that certain aspects of the nonnative structure can be maintained while others can fluctuate with the solvent conditions. The variable nature of the hydrophobic cluster could perhaps be important for burying these groups in the molten globule state. We have also shown that residues 106−111 can adopt nativelike secondary structure without disrupting the nonnative Y103−H107 interaction that is known to be present in the acid-induced molten globule state of αLA.

REFERENCES

1. Dyson H. J., and Wright, P. E. (1996) *Annu. Rev. Phys. Chem. 47*, 369−395.
2. Wüthrich, K. (1994) *Curr. Opin. Struct. Biol. 4*, 93−99.
3. Rothwarf, D. M., and Scheraga, H. A. (1996) *Biochemistry 35*, 13797−13807.
4. Kiefhaber, T., Grunert, H.-P., Hahn, U., and Schmid, F. X. (1992) *Proteins: Struct., Funct., Genet. 12*, 171−179.
5. Sosnick, T. R., Mayne, L., Hiller, R., and Englander, S. W. (1994) *Nat. Struct. Biol. 1*, 149−156.
6. Neri, D., Billeter, M., Wider, G., and Wüthrich, K. (1992) *Science 257*, 1559−1563.
7. Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., and Fenna, R. E. (1991) *J. Mol. Biol. 221*, 571−581.
8. Harata, K., and Muraki, M. (1992) *J. Biol. Chem. 267*, 1419−1421.
9. Schulman, B. A., and Kim, P. S. (1996) *Nat. Struct. Biol. 3*, 682−687.
10. Demarest, S. J., Fairman, R., and Raleigh, D. P. (1998) *J. Mol. Biol. 283*, 279−291.
11. Alexandrescu, A. T., Evans, P. A., Pitkeathly, M., Baum, J., and Dobson, C. M. (1993) *Biochemistry 32*, 1707−1718.
12. Smith, L. J., Alexandrescu, A. T., Pitkeathly, M., and Dobson, C. M. (1995) *Structure 2*, 703−712.
13. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
14. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
15. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
16. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.
17. Kim, Y., and Prestegard, J. H. (1989) *J. Magn. Reson. 84*, 9−13.
18. Griesinger, C., Sørenson, O. W., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.
19. Griesinger, C., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 261−271.
20. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gó, N., and Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
21. Brünger, A. T. (1992) *X-PLOR version 3.1. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
22. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) *Protein Eng. 2*, 27−38.
23. Nilges, M., Clore, G. M., and Gronenborn, A. M. *FEBS Lett. 239*, 129−136 (1988).
24. Nilges, M., Kuszewski, J., and Brünger, A. T. (1991) *Computational Aspects of the Study of Biological Macromolecules by NMR*, Plenum Press, New York.
25. Kuszewski, J., Nilges, M., and Brünger, A. T. (1992) *J. Biol. NMR 2*, 33−56.
26. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
27. Herzberg, O., and Moult, J. (1991) *Proteins: Struct., Funct., Genet. 11*, 223−229.
28. Bolin, K. A., Pitkeathly, M., Miranker, A., Smith, L. J., and Dobson, C. M. (1996) *J. Mol. Biol. 261*, 443−453.
29. Ramírez-Alvarado, M., Blanco, F. J., and Serrano, L. (1996) *Nat. Struct. Biol. 3*, 612.
30. Blandamer, M. J., Burgess, J., Cooney, A., Cowles, H. J., Horn, I. H., Martin, K. J., Morcom, K. W., and Warrick, P. (1990) *J. Chem. Soc., Faraday Trans. 86*, 2209−2213.
31. Cammers-Goodwin, A., Allen, T. J., Oslick, S. L., McClure, K. F., Lee, J. H., and Kemp, D. S. (1996) *J. Am. Chem. Soc. 118*, 3082−3090.
32. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI990320Z