BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 4707–4719 4707
DOI: 10.1021/bi802299j

The Unfolded State of the C-Terminal Domain of the Ribosomal Protein L9 Contains Both Native and Non-Native Structure†

Bing Shan,§ David Eliezer,*,‡ and Daniel P. Raleigh*,§,#

§*Department of Chemistry, State University of New York at Stony Brook, Stony Brook, New York 11794-3400,* ‡*Department of Biochemistry and Program in Structural Biology, Weill Cornell Medical College, New York, New York 10065, and* #*Graduate Program in Biochemistry and Structural Biology, and Graduate Program in Biophysics, State University of New York at Stony Brook, Stony Brook, New York 11794*

Received December 16, 2008. Revised Manuscript Received March 20, 2009

ABSTRACT: Interest in the structural and dynamic properties of unfolded proteins has increased in recent years owing to continued interest in protein folding and misfolding. Knowledge of the unfolded state under native conditions is particularly important for obtaining a complete picture of the protein folding process. The C-terminal domain of protein L9 is a globular α, β protein with an unusual mixed parallel and antiparallel β-strand structure. The folding kinetics and equilibrium unfolding of CTL9 strongly depend on pH, and follow a simple two state model. Both the native and the unfolded state can be significantly populated at pH 3.8 in the absence of denaturant, allowing the native state and the unfolded state to be characterized under identical conditions. Backbone ^{15}N, ^{13}C, ^{1}H and side-chain $^{13}C_{\beta}$, $^{1}H_{\beta}$ chemical shifts, amide proton NOEs, and ^{15}N R_2 relaxation rates were obtained for the two conformational states at pH 3.8. All the data indicate that the pH 3.8 native state is well folded and is similar to the native state at neutral pH. There is significant residual structure in the pH 3.8 unfolded state. The regions corresponding to the two native state α-helices show strong preference to populate helical φ and ψ angles. The segment that connects α-helix 2 and β-strand 2 has a significant tendency to form non-native α-helical structure. Comparison with the pH 2.0 unfolded state and the urea unfolded state indicates that the tendency to adopt both native and non-native helical structure is stronger at pH 3.8, demonstrating that the unfolded state of CTL9 under native-like conditions is more structured. The implications for the folding of CTL9 are discussed.

The past decade has witnessed increasing interest in the properties of the unfolded state ensemble of proteins, motivated in large part by the realization that unfolded states can contain significant amounts of nonrandom residual structure (*1*−*7*). Structural characterization of unfolded proteins is important because the unfolded state is the starting state for protein folding, the reference state for protein engineering studies, and can be the starting state for pathological protein aggregation. In addition, natively unfolded polypeptides play important roles in a variety of biological processes (*2*, *8*). The goal of the present study is to characterize the unfolded state and native state of the C-terminal domain of the ribosomal protein L9 (CTL9),[1] in the absence of denaturant under exactly the same solvent conditions, and to compare the structural propensities of this unfolded state to unfolded states populated under harsher conditions.

Proteins can be unfolded by high concentration of denaturants, extreme pH, temperature or high pressure. These states can be populated at equilibrium, but it is now widely recognized that the properties and structural propensities of unfolded proteins vary significantly depending on the experimental conditions (*9*−*12*). High concentrations of denaturants generally abolish residual secondary structure and lead to a more expanded unfolded state. In contrast, native-like or non-native structure has been found in a variety of unfolded proteins under less harsh conditions (*13*−*15*). Obviously, the unfolded state populated under physiological conditions is the most interesting

†This work was supported by NIH grants GM70941 (DPR) and AG019391 and AG025440 (DE) and by the Irma T. Hirschl Foundation and a gift from Herbert and Ann Siegel (DE). DPR and DE are members of the New York Structural Biology Center, which is a STAR center supported by the New York State Office of Science, Technology and Academic Research, by NIH grant P41 GM66354, and received funds from NIH, the Keck Foundation, New York State, and the NYC Economic Development Corporation for the purchase of 900 MHz spectrometers.

*Author to whom correspondence should be addressed. (D.P.R.) Telephone: 631-632-9547; fax: 631-632-7960; e-mail: draleigh@notes.cc.sunysb.edu. (D.E.) Telephone: 212-746-6557; fax: 212-746-4843; e-mail: dae2005@med.cornell.edu.

[1]CTL9, the C-terminal domain of the ribosomal protein L9; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt; PFG-NMR, pulse field gradient nuclear magnetic resonance; R_g, radius of gyration; R_h, radius of hydration; $\Delta\delta$, the difference between the chemical shift δ of a given residue in a protein and the random coil chemical shift.

Published on Web 03/20/2009

since it is most relevant to folding and is the proper thermodynamic reference state for protein engineering studies. Unfortunately, unfolded states under native conditions are extremely difficult to access experimentally because the free energy of folding favors the native state under physiological conditions and because folding is normally highly cooperative. Typically indirect approaches need to be applied to study the conformational propensities of the unfolded state under native conditions. These include the analysis of peptide fragments, studies of destabilized point mutants or truncation mutants, analysis of the pH dependence of stability, and H/D exchange measurements to name a few (*4*, *16*, *17*). While very useful, these approaches do not directly probe the conformational propensities of the unfolded state ensemble populated by the full-length wild type protein under conditions where the native state is stable. In a few favorable cases mild, near native conditions have been found where the folded state and unfolded state ensemble are both well populated in the absence of denaturant. Undoubtedly the best characterized such system is the drkN SH3 domain. This marginally stable protein populates the native state and unfolded state in an approximately 1:1 ratio, and the states are in slow exchange on the NMR time scale. Forman-Kay and coworkers have exploited this behavior to conduct extensive NMR studies of the drkN SH3 unfolded state (*15*, *18*−*21*). The picture that emerged from these studies is one of an unfolded state that is compact and in which certain regions of the polypeptide chain have a preference for native-like structure while other regions have preferences for non-native helical structure.

Unfolded state structure may play a role in folding by limiting the initial conformational search, a feature that has been proposed to be particularly important for some rapid folding proteins. Conversely significant non-native structure could slow folding if it needs to be broken as part of the rate limiting step in folding. Irrespective of its effect on folding, recent studies have clearly demonstrated that unfolded state interactions can be targeted to stabilize proteins (*4*, *5*, *22*). In addition, unfolded state interactions which are modulated by mutation can affect the interpretation of the commonly employed protein engineering approach for characterizing protein folding transition states (*23*). All of these factors have helped to motivate studies of the unfolded state ensemble.

Here we take advantage of the fact that CTL9 populates the folded and unfolded ensemble in the absence of denaturant at slightly acidic pH to characterize the unfolded ensemble under near native conditions. CTL9 is a 92 residue globular protein derived from the ribosomal protein L9 (Figure 1). The protein adopts an interesting fold made up of two α-helices and a three stranded mixed parallel, antiparallel β-sheet. The protein contains three histidines, making the equilibrium unfolding strongly pH dependent, and the domain undergoes reversible unfolding as the pH changes between 7 and 2 (*24*, *25*). At the middle point of this transition, which is around pH 3.8, both the native state and the unfolded state are populated significantly and are in slow exchange on the NMR time scale, allowing the structural properties of both states to be measured under identical conditions. Previous work has

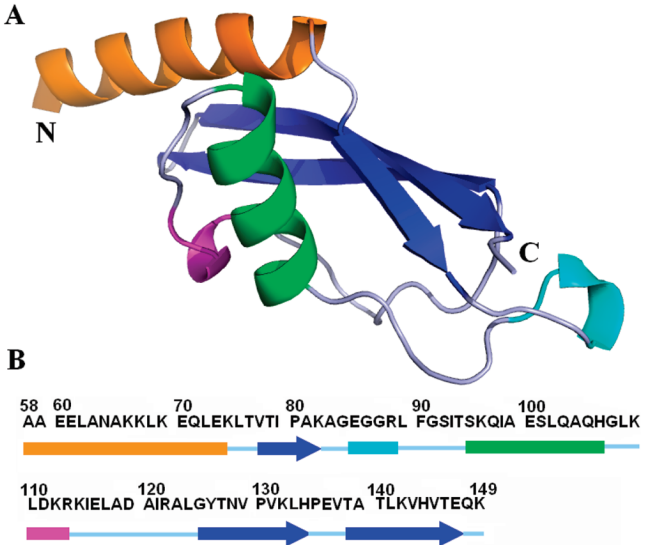

FIGURE 1: Structure of CTL9. (A) Ribbon diagram of CTL9 (residues 58−149 of protein L9). Protein data bank entry 1DIV. The N- and C-termini are labeled. (B) The primary sequence of CTL9 together with a schematic representation of the secondary structural elements (arrows represent β-strands and colored cylinders α-helices and 3_{10}-helices). The ribbon diagram was made using PyMol.

shown that the unfolded state ensemble populated under near native conditions is compact with a radius of hydration of 25.1 Å. In contrast, under strongly denaturing conditions (pH 3.8, 8 M urea) the denatured ensemble is expanded with a radius of hydration of 33.6 Å (*26*). The CD spectrum of the pH 3.8 unfolded state indicates the presence of residual helical structure but obviously could not localize it to specific regions of the chain.

MATERIALS AND METHODS

Protein Expression and Purification. ^{15}N-Labeled CTL9 and ^{15}N, ^{13}C-labeled CTL9 were overexpressed in M9 minimal medium and purified as previously described (*14*). The purity was confirmed using reverse phase HPLC, and the identity of the protein was validated using MALDI mass spectroscopy. The yield of pure protein was 27 mg/L of media.

NMR Experiments. Protein samples for NMR experiments were dissolved in 90% H_2O/10% D_2O at pH 3.8, with 20 mM sodium acetate and 100 mM NaCl. The protein concentration was about 1 mM. All NMR experiments were carried out on an 800 MHz Bruker spectrometer equipped with a cryo probe at the New York Structural Biology Center. All the spectra were collected at 25 °C. In all NMR experiments, the ^{1}H dimension was centered at the water resonance and the offset of the ^{15}N dimension was set at 118.0 ppm.

^{15}N−^{1}H correlated heteronuclear single-coherence (HSQC) spectrum (*27*) was acquired using 1024 × 256 complex points with eight scans per increment. The spectral widths were 9615.4 Hz for the ^{1}H dimension and 2433.1 Hz for the ^{15}N dimension. Water suppression was accomplished using the watergate sequence. The following triple resonance experiments were performed to establish backbone connectivities, on a ^{15}N, ^{13}C-labeled protein sample: HNCO, HNCACO, HNCACB, and CBCACONH. Typical spectral widths were 8389.3 Hz

for the ^{1}H dimension and 2270.7 for the ^{15}N dimension. Constant time HNCO (*28*, *29*) and constant time HNCACO (*29*, *30*) experiments were acquired with 1024 × 80 × 144 complex points. The ^{13}CO dimension has a spectral width of 2817.7 Hz and centered at 176 ppm. Constant time HNCACB (*31*, *32*) and constant time CBCACONH (*32*, *33*) experiments comprised 1024 × 80 × 160 complex points. The ^{13}C dimension has a spectral width of 14084.5 Hz, with the ^{13}C offset set at 39 ppm. In order to resolve the ambiguities caused by the severe resonance overlap, a set of $^{15}N-^{1}H$ HSQC spectra were collected between pH 2.0 and pH 5.5 with 0.4 pH unit increments. By tracking peaks, it is either possible to directly transfer the assignments previously determined for the pH 2.0 unfolded state and the pH 5.5 native state or constrain the possible identity of the resonances of the pH 3.8 spectrum. To determine the $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ chemical shifts, a three-dimensional (3D) constant time HBHACONH experiment (*32*, *33*) was acquired with a data matrix of 1024 (direct ^{1}H) × 80 (^{15}N) × 160 (indirect ^{1}H) complex points. The indirect ^{1}H dimension has a spectral width of 7204.6 Hz. In all the above triple resonance experiments, the watergate sequence was used to suppress the water signal.

3D ^{15}N-edited HSQC-NOESY (*34*) and the ^{15}N HSQC-NOESY-HSQC (*35*) experiments were performed on a ^{15}N-labeled sample of CTL9. The ^{15}N HSQC-NOESY spectrum was acquired with spectral widths of 9615 (HN), 2433 (^{15}N), and 9602 Hz (^{1}H) in a data matrix of 1024 × 90 × 256 complex points. The ^{15}N HSQC-NOESY-HSQC spectrum was recorded with spectral widths of 8389 (NH), 2190 (^{15}N) and 2190 (^{15}N) Hz and 1024 × 140 × 170 complex points in the t3 (NH), t2 (^{15}N), and t1 (^{15}N) dimensions, respectively. Coherence selection with gradients and sensitivity enhancement were used in both experiments. Water suppression was achieved using a water flip back pulse. A mixing time of 300 ms was used to acquire both NOESY spectra. To cancel out possible relaxation effects on the NOE intensities, the NOE peak volumes are normalized as the ratio of the *NN*(i,i + 1) cross peak to the *NN*(i, i) diagonal peaks.

Relaxation experiments were performed on a ^{15}N-labeled sample, using methods described previously (*36*). The spectra were recorded at 10 delay times: 16.96, 33.92, 50.88, 67.84, 84.8, 101.76, 118.72, 135.68, 152.64, and 169.6 ms, with repetition at 50.88, 101.76, and 152.64 ms to allow the estimation of the uncertainty. Each spectrum was collected with four scans using 1024 × 256 complex points. The spectral widths were 8808.2 Hz (^{1}H) and 2433.1 (^{15}N). A recycle delay of 3 s was used.

All spectra were processed using NMRPipe software (*37*), and chemical shift assignments were made using NMRView J (*38*). NOESY peak volumes were measured using NMRViewJ. All the chemical shifts were referenced to the absolute frequency of DSS at 0.00 ppm. Sequence-dependent corrections were made using the protocol reported by Schwarzinger and co-workers (*39*). The random coil values of Wishart were used to calculate the secondary chemical shifts (*40*). SSP analysis was performed using the software provided by Professor Forman-Kay on her Web site (http://pound.med.utoronto.ca/software.html). R_2 relaxation rates were determined using the automated program in NMRView J by fitting the peak intensities to eq 1 which describes a two-parameter exponential decay:

$$I(t) = I_0 \exp(-t/T_2) \quad (1)$$

where $I(t)$ is the peak intensity after a delay of time t and I_0 is the intensity at the time zero. The R_2 rates were analyzed using a simple model by fitting the experimental R_2 rates to eq 2:

$$R_2(i) = R_2(\text{int}) \sum_{j=1}^{N} \exp\left(-\frac{|i-j|}{\lambda}\right) \quad (2)$$

where $R_2(i)$ is the experimental R_2 value for residue i, R_2(int) is the intrinsic relaxation rate which depends on the temperature and viscosity of the solvent, λ is the apparent persistence length of the chain, and N is the total number of residues in the protein (*3*).

RESULTS

Sequence-Specific Assignment of the Native State and the Unfolded State of CTL9 at pH 3.8. CTL9 is well folded at pHs above 5, and is fully unfolded at pH 2.0 (*14*, *25*). At pH 3.8, both the native state and the unfolded state are significantly populated. The two states are in slow exchange on the NMR chemical shift time scale, allowing resonances from both states to be observed. The $^{15}N-^{1}H$ HSQC spectrum of CTL9 at pH 3.8 is displayed in Figure 2. The peaks are sharp, consistent with both states being monomeric. A number of peaks mostly from the native state are clearly resolved, but the rest of the native state and most of the unfolded state resonances are clustered in a limited frequency range at the center of the spectrum. The limited dispersion is typical for an unfolded protein but is worse than often observed due to the overlap with the native state peaks and because of the small number of aromatic residues in CTL9. Nevertheless, virtually complete assignments could be obtained for both states. A set of NMR experiments including the HNCACB, CBCACONH, HNCO, and HNCACO experiments were conducted to determine inter-residue backbone connectivities and to obtain backbone assignments. In addition, a set of HSQC spectra were collected over the pH range of 2−5.5 to resolve ambiguous assignments. The HBHACONH experiment was used to obtain $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ assignments. Complete ^{15}N, ^{13}C, ^{1}H backbone assignments as well as $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments were obtained for 80 residues in the native state and 86 residues of the unfolded state. Partial assignments were obtained for additional eight residues of the native state and five residues of the unfolded state. Overall at least partial assignments were obtained for every nonproline residue in the pH 3.8 unfolded state. All of the assignments have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/).

Analysis of the ^{13}C and ^{1}H Chemical Shifts Indicates Significant Native and Non-Native Secondary Structure in the Unfolded State. It is well-known that chemical shifts are very sensitive to protein structure. The deviations of the observed chemical shifts from random coil values,

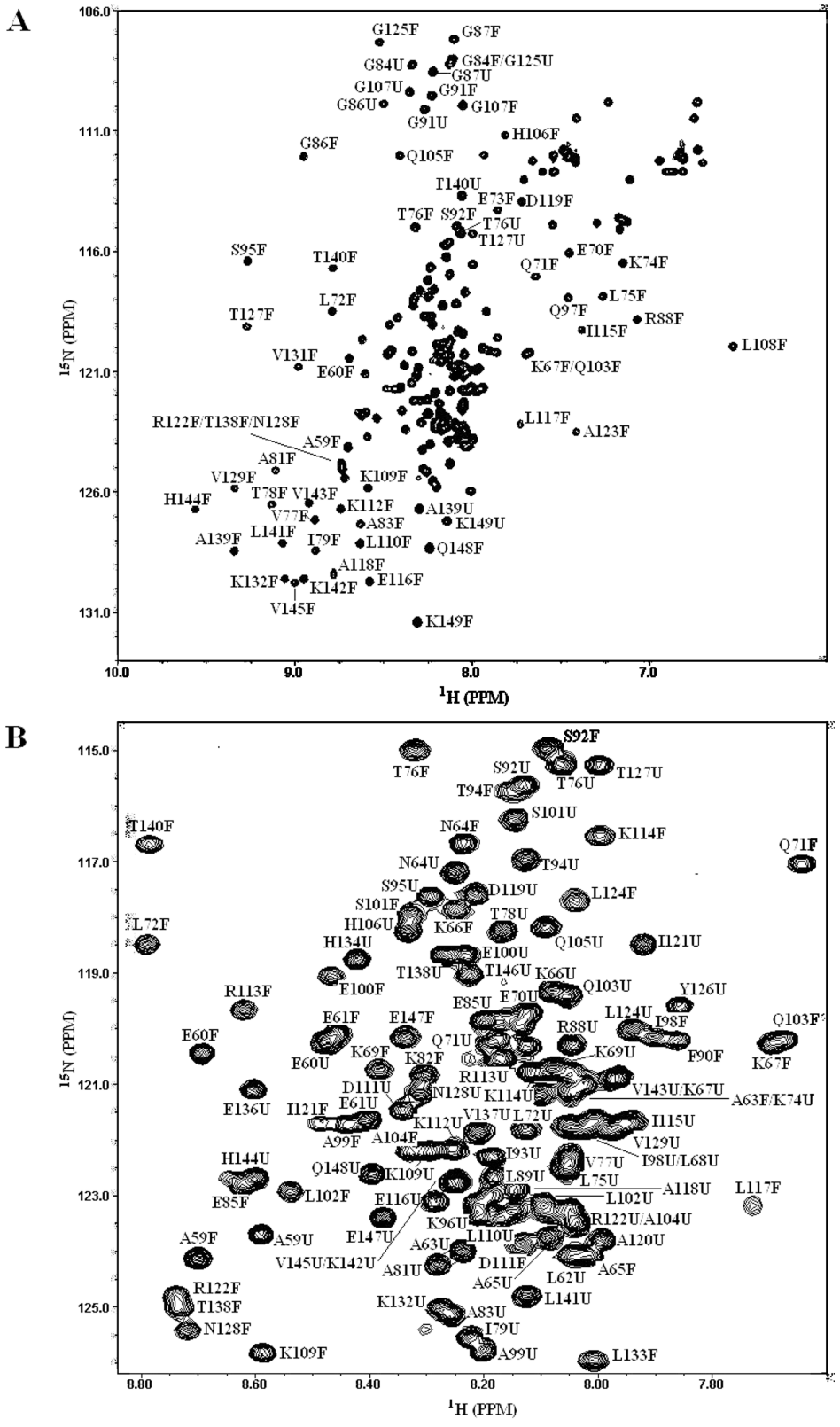

FIGURE 2: ^{15}N-HSQC spectrum (panel A) of CTL9 with assignments of peaks indicated. U denotes unfolded state resonances and F folded state peaks. The spectrum was recorded at pH 3.8 and 25 °C. Panel B is an expansion of the central region of the spectrum.

so-called secondary chemical shifts, are very accurate predictors of protein secondary structure (*41*−*43*). We used the chemical shift data reported by Wishart (*40*) as random coil values to determine the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, and ^{13}CO secondary shifts of both the native state and the unfolded state of CTL9 at pH 3.8. The secondary shifts were compared to those of the pH 2.0 unfolded state.

Figure 3 displays the $^{1}H_{\alpha}$ secondary shifts for CTL9. Negative $^{1}H_{\alpha}$ secondary shifts suggest α-helical structure and positive $^{1}H_{\alpha}$ secondary shifts are observed for β-sheet structure (*42*). Figure 3 compares the $^{1}H_{\alpha}$ secondary shifts for the pH 3.8 native state and the pH 3.8 unfolded state. The pH 3.8 native state shows the largest $^{1}H_{\alpha}$ secondary shifts, with values ranging from −0.73

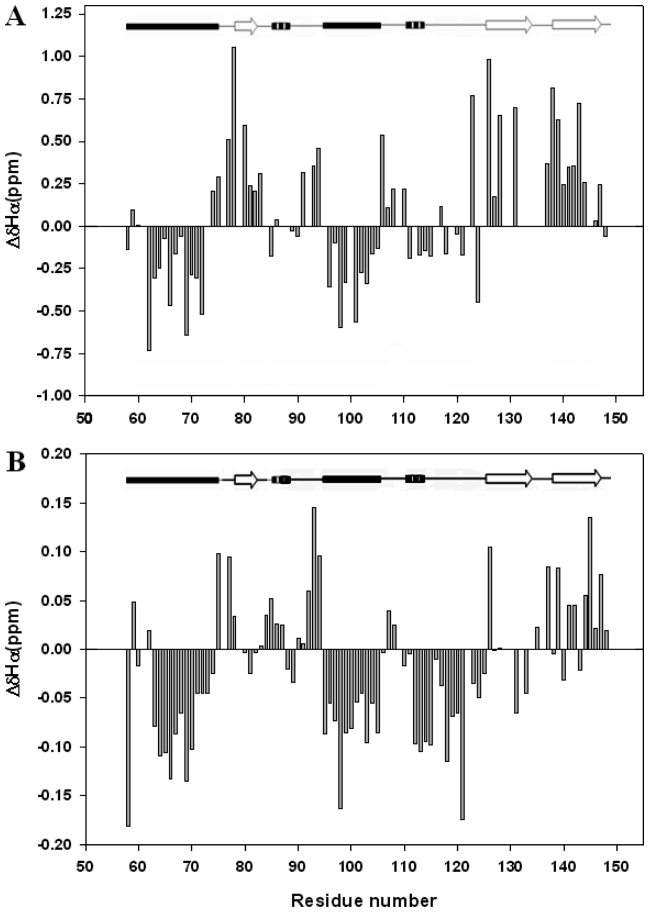

FIGURE 3: Plots of secondary $^1H_\alpha$ shifts as a function of residue number. Random coil values from Wishart (*40*) were used together with sequence-specific corrections (*39*). (A) $^1H_\alpha$ secondary shifts for the pH 3.8 native state, (B) $^1H_\alpha$ secondary shifts for the pH 3.8 unfolded state. A schematic representation of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel (arrows represent β-strands and filled cylinders α-helices, dashed cylinders 3_{10}-helices, and single lines loop regions).

ppm to 1.05 ppm. The pattern of secondary shifts agrees well with the native structure of CTL9 at pH 5.5. As expected, the segments corresponding to the two α-helices in the native state have negative secondary $^1H_\alpha$ shifts, with an average of −0.17 ppm for α-helix 1 (residues 58−74) and −0.18 for α-helix 2 (residues 95−106). Positive $^1H_\alpha$ secondary shifts are detected from regions that form the β-strands in the native state, namely, residues 79−82, residues 125−134, and residues 138−147. The average deviations are 0.52 ppm for strand 1, 0.63 ppm for strand 2, and 0.35 ppm for strand 3. The two 3_{10} helix regions (residues 85−88 and 110−113) exhibit small deviations with no clear trend. In comparison, the pH 3.8 unfolded state exhibits smaller $^1H_\alpha$ secondary shifts, as expected for an unfolded protein. However, the overall pattern is similar to that observed for the native state. In particular, the residues that comprise the two α-helices display contiguous negative $^1H_\alpha$ secondary shifts. The average deviation is −0.07 ppm for both helices. The segment of the polypeptide chain, which connects α-helix 2 and β-strand 2, residues 107−124, also displays negative $^1H_\alpha$ secondary shifts, −0.06 ppm on average. The deviations are very small for the first two β-strands with no obvious trends, while the third strand exhibits generally positive

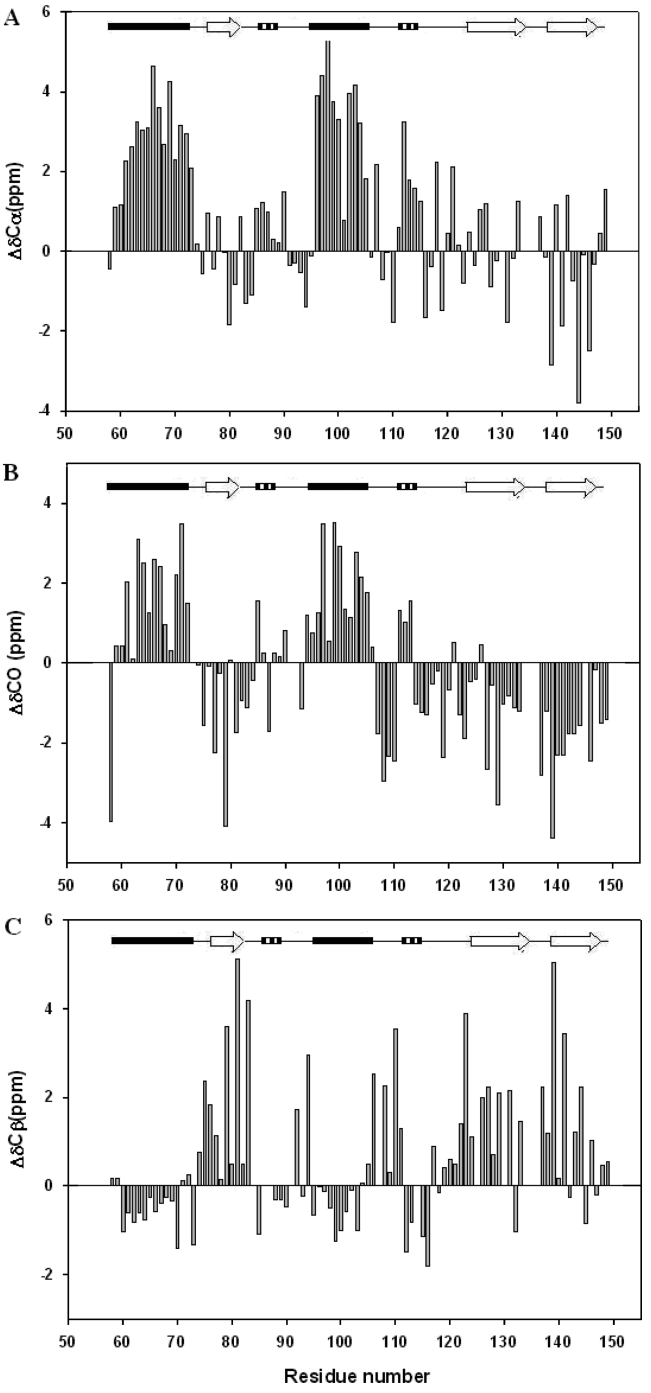

FIGURE 4: Plots of secondary ^{13}C shifts for the pH 3.8 native state. (A) $^{13}C_\alpha$ secondary shifts, (B) carbonyl ^{13}CO secondary shifts, (C) $^{13}C_\beta$ secondary shifts. Sequence-dependent corrections were made. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

deviations (0.04 ppm on average). The $^1H_\alpha$ secondary shifts data clearly suggest a propensity to populate helical ψ and φ values in the pH 3.8 unfolded state for the residues which are α-helical in the folded state. $^1H_\alpha$ secondary shifts are very useful, but because of the relatively small dispersion in ^1H chemical shifts, more accurate and reliable information can be obtained by the analysis of ^{13}C secondary shifts.

$^{13}C_\alpha$ and ^{13}CO resonances are shifted downfield in α-helices and upfield in β-strands. The trend is reversed for $^{13}C_\beta$ secondary shifts (*40*, *44*). The $^{13}C_\alpha$, $^{13}C_\beta$, and ^{13}CO

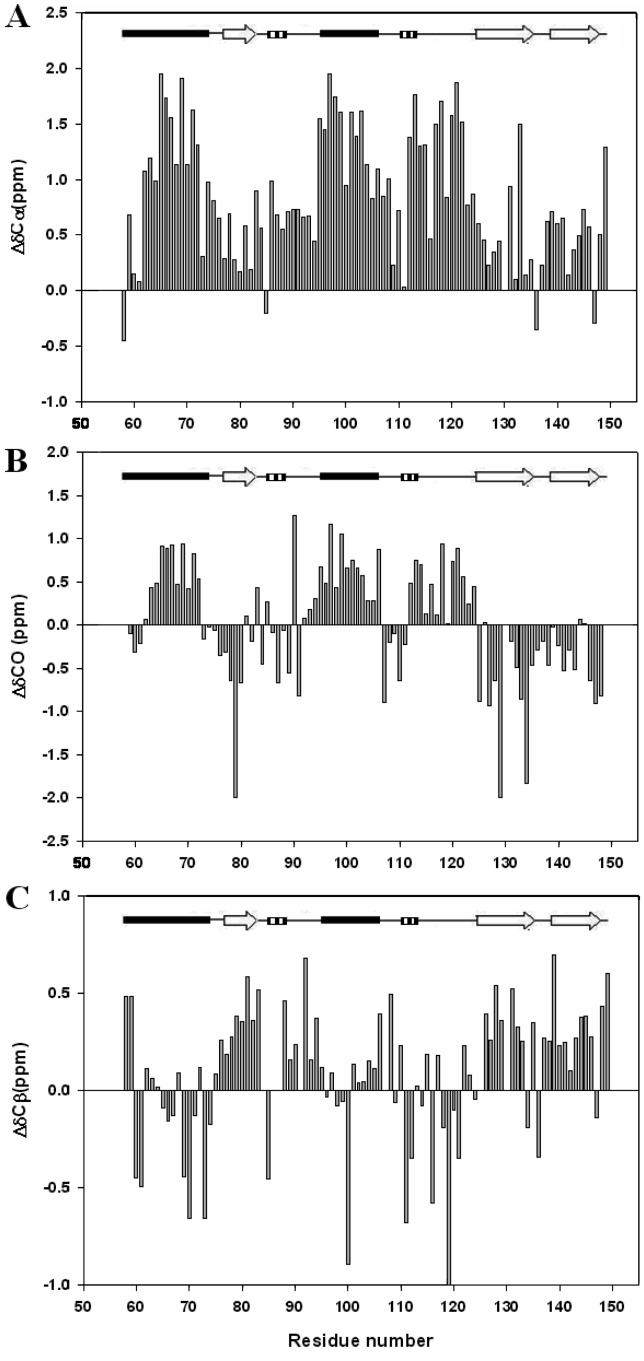

FIGURE 5: Plots of secondary ^{13}C shifts for the pH 3.8 unfolded state. (A) $^{13}C_\alpha$ secondary shifts, (B) carbonyl ^{13}CO secondary shifts, (C) $^{13}C_\beta$ secondary shifts. Sequence-dependent corrections were made for the ^{13}CO shifts. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

secondary shifts of the pH 3.8 native state and the pH 3.8 unfolded state are plotted in Figures 4 and 5. Table 1 summarizes the average secondary shifts for stretches of the primary sequence which display consecutive large deviations from zero. The overall pattern of $^{13}C_\alpha$ shifts is consistent with the $^1H_\alpha$ secondary shifts. Continuous positive deviations are detected for the pH 3.8 native state (Figure 4A), in the two α-helical regions. The average value is 2.46 ppm for α-helix 1 and 2.87 ppm for α-helix 2. Positive deviations are also found in residues 110−114, which corresponds to the second 3_{10} helix in the crystal

Table 1: Average Secondary Shifts and SSP Values in Secondary Structural Elements of CTL9

	pH 3.8 native state					pH 3.8 unfolded state				
	$^1H_\alpha$	$^{13}C_\alpha$	$^{13}C_\beta$	^{13}CO	SSP	$^1H_\alpha$	$^{13}C_\alpha$	$^{13}C_\beta$	^{13}CO	SSP
α-helix 1	−0.17	2.46	−0.40	1.14	0.86	−0.07	1.02	−0.12	0.38	0.46
α-helix 2	−0.18	2.87	−0.42	1.84	0.82	−0.07	1.41	0.00	0.66	0.41
loop	−0.01	0.51	0.64	−0.89	0.16	−0.06	1.09	−0.15	0.25	0.38
strand 1	0.52	−0.20	1.97	−1.39	−0.53	0.02	0.38	0.39	−0.68	0.11
strand 2	0.63	0.01	1.37	−1.21	−0.40	−0.01	0.53	0.31	−0.87	0.21
strand 3	0.35	−1.07	1.31	−2.09	−0.19	0.04	0.44	0.27	−0.34	0.13

structure. β-strand 3 exhibits negative deviations with an average of −1.07 ppm, while oscillating deviations with no clear trend are found for β-strands 1 and 2 (averages of −0.20 and 0.01 ppm, respectively). The pH 3.8 unfolded state displays $^{13}C_\alpha$ secondary shifts (Figure 5A) which are smaller than those of the native state (Figure 4A), but the values are clearly significant. The largest deviations are observed in the regions that correspond to the two α-helices in the native state. The average deviation is 1.02 ppm for α-helix 1 and 1.41 ppm for α-helix 2. Both of these values are significantly larger than the values previously observed for the pH 2.0 unfolded state, 0.32 ppm on average for α-helix 1 and 0.62 ppm on average for α-helix 2 (14). The segment linking α-helix 2 and β-strand 2 (residues 107−124) also shows secondary $^{13}C_\alpha$ shifts in the pH 3.8 unfolded state (1.09 ppm) that are larger than those in the pH 2.0 unfolded state (0.55 ppm). Small positive deviations are detected in the β-strand regions. The average is 0.38 ppm for β-strand 1, 0.53 ppm for β-strand 2, and 0.44 ppm for β-strand 3.

The ^{13}CO secondary shifts display a broadly similar pattern to the $^{13}C_\alpha$ secondary shifts. In the pH 3.8 native state, large positive ^{13}CO secondary shifts are detected in the two α-helical regions, while negative deviations are found in the β-strand regions (Figure 4B). The average is 1.14 ppm for α-helix 1 and 1.84 ppm for α-helix 2. The segment linking α-helix 2 and β-strand 2 displays an average deviation of −0.89 ppm, and the values are −1.39 ppm, −1.21 ppm, and −2.09 ppm for the three β-strands, respectively. The ^{13}CO secondary shifts of the pH 3.8 unfolded state are generally of smaller amplitude than those of the pH 3.8 native state but are significant (Figure 5B). The averages are 0.38 and 0.66 ppm for α-helix 1 and α-helix 2, respectively. The segment linking α-helix 2 and β-strand 2 has positive deviations with an average of 0.25 ppm. This is an interesting observation since the region displays negative deviations in the pH 3.8 folded state. In the folded state, this region contains a short 3_{10} helix between residues 110 to 113, while the rest of the segment does not adopt any classical secondary structure. The averages for the three β-strands are −0.68 ppm, −0.87 ppm, and −0.34 ppm, respectively. Again, the deviations are larger than those reported for the pH 2.0 unfolded state.

The pattern of $^{13}C_\beta$ secondary shifts is also consistent with the $^1H_\alpha$, $^{13}C_\alpha$, and ^{13}CO shifts. Consecutive negative $^{13}C_\beta$ secondary shifts are observed in the two α-helical regions in the pH 3.8 native state (−0.40 ppm on average

for α-helix 1 and −0.42 ppm on average for α-helix 2), while positive $^{13}C_\beta$ secondary shifts are detected in the β-strand regions (Figure 4C). The average values for the three β-strands are 1.97 ppm, 1.37 ppm, and 1.31 ppm, respectively, while the segment that connects α-helix 2 and β-strand 2 has an average deviation of 0.63 ppm. A largely similar pattern of $^{13}C_\beta$ secondary shifts are found for the pH 3.8 unfolded state, however once again, the magnitudes of the deviations are smaller. The average is −0.12 ppm for α-helix 1, 0 ppm for α-helix 2, and −0.15 ppm for the segment connecting α-helix 2 and β-strand 2. Notice that the sign of the deviations for this segment are reversed relative to the pH 3.8 folded state. This provides additional evidence that this region of the polypeptide chain has a propensity to populate non-native structure in the pH 3.8 unfolded state. The averages for the three β-strands are 0.39 ppm, 0.31 ppm, and 0.27 ppm, respectively.

Analysis of Chemical Shift Differences and Secondary Structural Propensity Scores Indicates Significant Helical Structure in the Unfolded State. Combinations of secondary shifts often provide a more consistent picture of secondary structural propensities and can avoid problems with chemical shift referencing. A common approach is to calculate the difference between the $^{13}C_\alpha$ secondary shifts and the $^{13}C_\beta$ secondary shifts, $\Delta\delta C_\alpha - \Delta\delta C_\beta$ values. Positive $\Delta\delta C_\alpha - \Delta\delta C_\beta$ values indicate α-helical structure and negative values suggest a propensity to form β-strand structure (*45*). Figure 6A plots the $\Delta\delta C_\alpha - \Delta\delta C_\beta$ values for the pH 3.8 native state. The largest positive values are found in the two α-helix regions, 2.86 and 3.29 ppm on average. Negative values are detected in the three β-strands, with averages of −2.17 ppm, −1.36 ppm, and −2.38 ppm, respectively. The value for the region connecting α-helix 2 and β-strand 2 is −0.08 ppm. The $\Delta\delta C_\alpha - \Delta\delta C_\beta$ values for the pH 3.8 unfolded state reveal a tendency to form helical structure in certain regions (Figure 6B). The average is 1.24 ppm for α-helix 1 and 1.41 ppm for α-helix 2, while no obvious trends are detected in the three β-strands. Interestingly, residues 107−124 display positive $\Delta\delta C_\alpha - \Delta\delta C_\beta$ values in the pH 3.8 unfolded state but not in the folded state. In fact, the average deviation in this region is 1.25 ppm in the pH 3.8 unfolded state, which is comparable to the values detected for α-helix 1 and α-helix 2. Again, this is consistent with a tendency to form non-native structure. This region also displayed positive deviations in the previously characterized pH 2.0 unfolded state although they were significantly smaller (*14*).

Recently, the secondary structure propensity (SSP) analysis developed by Forman-Kay and co-workers has been used to analyze secondary structure propensity in natively unfolded proteins (*45*). The SSP method uses chemical shifts of different nuclei to calculate a single score to represent the propensity of a given residue to populate secondary structure. A SSP score of 1 is taken to mean a well-formed α-helix and a value of −1 to represent a β-strand. Figure 7 displays the results of the SSP analysis applied to the pH 3.8 native state and the pH 3.8 unfolded state of CTL9. The data are summarized in Table 1. In the folded state, the residues comprising α-helix 1 have SSP scores ranging from 0.59 to 1.03 with an average of 0.86. The values for α-helix 2 range from 0.54 to 1.10 with an average of 0.82. The region corresponding to β-strand 1 has SSP scores ranging

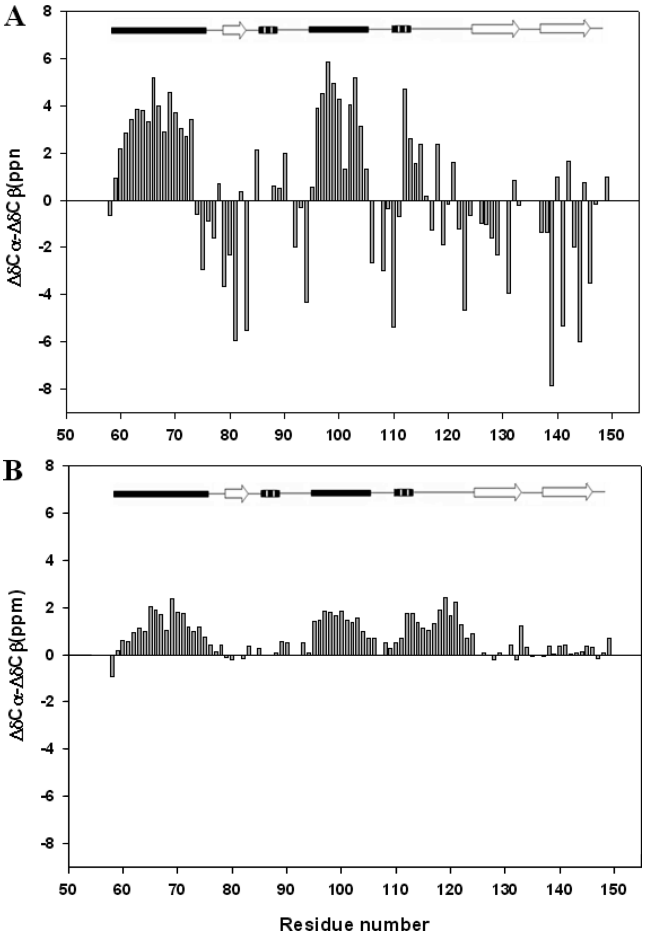

FIGURE 6: Plots of the difference in the $^{13}C_\alpha$ secondary shifts and the $^{13}C_\beta$ secondary shifts for (A) the pH 3.8 native state, (B) the pH 3.8 unfolded state. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

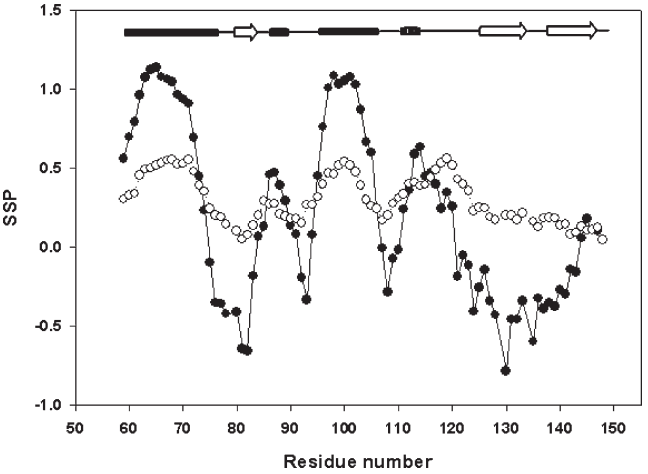

FIGURE 7: SSP analysis of the pH 3.8 native state (●) and the pH 3.8 unfolded state (○) of CTL9. Positive SSP values indicate a propensity to populate the helical region of the Ramachandran plot, while negative values indicate a preference for the β-strand region. $^{13}C_\alpha$, $^{13}C_\beta$, and $^{1}H_\alpha$ chemical shifts were used in the analysis. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

from −0.41 to −0.63 with an average of −0.53. The values for β-strand 2 range from −0.14 to −0.78 with an average of −0.40, and β-strand 3 has SSP scores ranging from −0.38 to 0.11 with an average of −0.11. The SSP scores for the pH 3.8 unfolded state indicate that the residues within helix-1 and helix-2 have a significant tendency to populate helical φ and ψ angles. In the pH 3.8 unfolded state, the region corresponding to α-helix 1 has SSP values ranging from 0.37 to 0.55 with an average of 0.46. For α-helix 2, the average is 0.41 with individual values ranging from 0.30 to 0.52. We have previously analyzed the pH 2.0 unfolded state and the helical regions have SSP scores of 0.28 ± 0.07 and 0.32 ± 0.07, indicating a weaker but still significant tendency to form helical structure. Clear differences between the folded and unfolded states are observed for the region linking α-helix 2 and β-strand 2, particularly for residues 107−110 and 118−124 where the average deviation is 0.39 between the pH 3.8 folded and unfolded states. Residues 111−117, which include the 3_{10} helix, display smaller differences with an average deviation of 0.04. The average SSP score for the entire region encompassing residues 107− 124 is 0.38 in the pH 3.8 unfolded state, and the individual values cover the range of 0.27−0.49. The same region in the native state, however, has a much smaller average SSP value, which is 0.16, and the individual values span the range from −0.15 to 0.47. In summary, the SSP plot clearly reveals that the residues which populate helical φ and ψ angles in the native state have significant propensity to do so in the pH 3.8 unfolded state, while the segment linking α-helix 2 and β-strand 2 has significant propensity to form non-native α-helical structure in the pH 3.8 unfolded state.

The SSP scores are correlated for the pH 3.8 native state and the pH 3.8 unfolded state, but in contrast, there is no significant correlation between the SSP scores for the pH 3.8 unfolded state and the low pH urea unfolded state. Plots comparing the respective correlations are shown in Figure 8. The correlation between the two pH 3.8 states is better for residues in the helix-1 and helix-2 regions than is observed for the full protein ($r^2 = 0.68$ vs $r^2 = 0.62$). No correlation is detected when comparing the SSP scores of these regions between the pH 3.8 unfolded state and the low pH urea unfolded state. This simple analysis reinforces the picture that the pH 3.8 unfolded state, that is, the unfolded state populated in the absence of denaturant, has significant tendency to preferentially adopt native-like φ, ψ angles in these regions. Note that the SSP scores of the loop region which was identified as populating non-native structure in the pH 3.8 unfolded state exhibits a significantly worse correlation with the native state with a r^2 of only 0.35; this is entirely consistent with the other experiments that revealed non-native structure in this region.

Amide Proton NOEs Confirm the Presence of Helical Structure in the Unfolded State. NOEs are powerful constraints for protein structure determination, providing long-range distance constraints on global topology and short-range constraints on local structure, including secondary structure. However, measurement of NOEs in unfolded proteins is difficult due to the highly fluctuating nature of unfolded state ensembles and often limited spectral resolution. Nevertheless, characteristic NOEs can be observed in unfolded proteins if the population of a conformation with defined secondary structure is significant and the resolution is adequate. Sequential $d_{NN}(i,i+1)$ NOEs are particularly useful for detecting helical structure in unfolded proteins and can be used to distinguish helical from β-strand or extended structures. Using the ^{15}N HSQC-NOESY experiment, sequential $d_{NN}(i,i+1)$ NOEs and $d_{\alpha N}(i,i+1)$ NOEs were detected throughout the primary sequence of CTL9 in the pH 3.8 native state. However, only 18 $d_{NN}(i,i+1)$ NOEs and 25 $d_{\alpha N}(i,i+1)$ NOEs could be identified in the pH 3.8 unfolded state using this experiment, due to the overlap of resonances and the high degeneracy of the proton chemical shifts. In contrast, the ^{15}N HSQC-NOESY-HSQC experiment disperses the amide resonances into two nitrogen dimensions, which have much higher chemical shift dispersion than the proton dimension. As a result, this experiment is typically employed to measure $d_{NN}(i,i+1)$ NOEs in unfolded proteins even though it suffers from sensitivity loss. Figure 9 shows a diagram of the observed $d_{NN}(i,i+1)$ NOEs in the pH 3.8 native state and the pH 3.8 unfolded state. Only 25 $d_{NN}(i,i+1)$ NOEs are detected in the pH 3.8 native state because of the low sensitivity of the ^{15}N HSQC-NOESY-HSQC experiment. Only helical regions displayed consecutive sets of $d_{NN}(i,i+1)$ NOEs. Thirty $d_{NN}(i,i+1)$ NOEs are observed in the pH 3.8 unfolded state. These include eight $d_{NN}(i,i+1)$ NOEs in residues 58−74 (α-helix 1), seven NOEs in residues 95−106 (α-helix 2), and seven NOEs in residues 107−124, which comprise the segment linking α-helix 2 and β-strand 2 in the native fold. Five observed NOEs within residues 107−124 result from residue pairs which have interamide proton distances greater than 4 Å in the X-ray structure. Two of the seven result from pairs that have interproton distances less than 3 Å, and both of these are in the regions of 3_{10} helix. The former set is consistent with the segment adopting non-native structure in the unfolded state. The only other continuous $d_{NN}(i,i+1)$ NOEs are found for residues G86 and G87, which are located in the first 3_{10} helix region in the native state. The $d_{NN}(i,i+1)$ NOEs are much less intense in the pH 3.8 unfolded state than in the pH 3.8 native state, suggesting that the residual structure in the pH 3.8 unfolded state is dynamic and populated by a subset of the ensemble.

^{15}N R_2 Relaxation Analysis of the pH 3.8 Native State and the pH 3.8 Unfolded State. Measurement of relaxation rates can provide structural and dynamic information about unfolded proteins (*3*, *46*). ^{15}N R_2 relaxation rates are particularly useful in the detection of deviations from random coil behavior in unfolded proteins. For a completely unstructured polypeptide, the ^{15}N R_2 rates are expected to be uniform in the middle of the polymer chain and slightly smaller at the two termini. Significantly large R_2 rates are usually assumed to reflect local structure or persistent hydrophobic clusters (*3*, *46*). Figure 10 plots the measured relaxation rates and relaxation rates calculated using the phenomenological random coil model developed by Schwalbe and co-workers (*46*). As expected, the native state exhibits large systematic positive deviations from the values predicted by the phenomenological random coil model (Figure 10A). The deviations observed

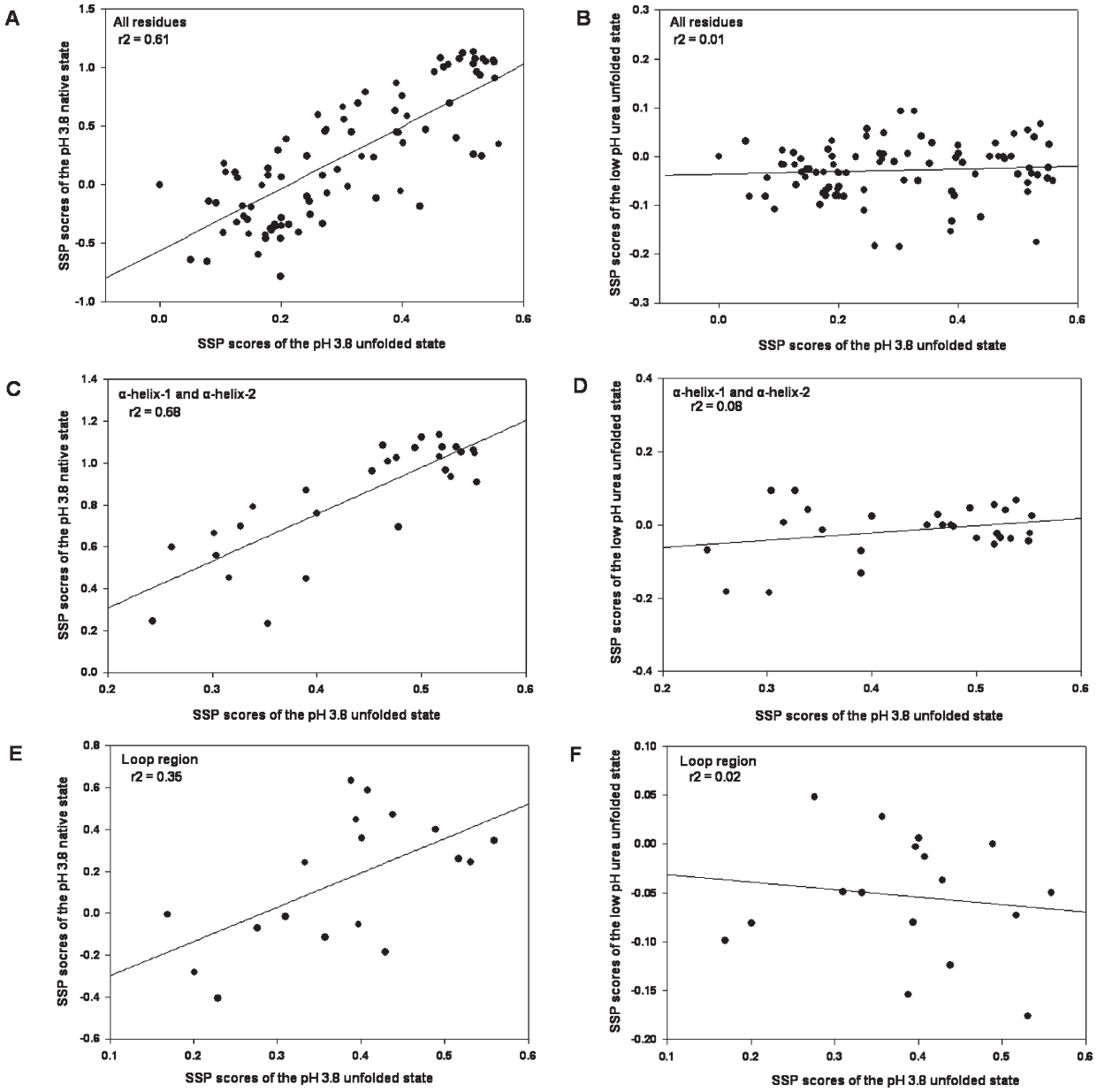

FIGURE 8: Comparison of the SSP scores of the pH 3.8 unfolded state with the SSP scores of the pH 3.8 native state and the low pH urea unfolded state. (A) Comparison of the pH 3.8 unfolded state with the pH 3.8 native state for all residues. (B) Comparison of the pH 3.8 unfolded state with the low pH unfolded state for all residues. (C) Comparison of the pH 3.8 unfolded state with the pH 3.8 native state for residues in α-helix 1 and α-helix 2. (D) Comparison of the pH 3.8 unfolded state with the low pH urea unfolded state for residues in α-helix 1 and α-helix 2. (E) Comparison of the pH 3.8 unfolded state with the pH 3.8 native state for residues in the loop region only. (F) Comparison of the pH 3.8 unfolded state with the low pH urea unfolded state for residues in the loop region only.

for the pH 3.8 unfolded state are much smaller but are clearly not zero and are considerably larger than the uncertainty in the data points (Figure 10B). The calculated R_2 values follow the expected inverted U-shaped curve but with some noticeable deviations. Regions that show the largest deviations include α-helix 2 and the segment between α-helix 2 and β-strand 2. Smaller but noticeable deviations are observed in α-helix 1. Interestingly, these are the regions that were identified as having the highest propensity to populate helical secondary structure in the unfolded ensemble by both the chemical shift analysis and by the NOE measurements. The fit to the phenomenological model yields (eq 2) a parameter λ which has been related to the apparent persistence length of the chain. The value obtained here is eight residues which is consistent with previous estimates (*3*, *46*). We note however that as rigorously defined it is not clear that the concept of persistence length is applicable to proteins.

DISCUSSION

We have obtained nearly complete backbone ^{15}N, ^{13}C, ^{1}H and side-chain $^{13}C_{\beta}$, $^{1}H_{\beta}$ chemical shifts for the pH 3.8 native state of CTL9 and the pH 3.8 unfolded state populated in the absence of denaturant. Direct measurement of the chemical shifts, NOEs, and R_2 relaxation rates of the two states allows the comparison of the structural and dynamic properties under identical native-like conditions. The secondary chemical shifts, the difference of $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ shifts, SSP analysis, the ^{15}N R_2 values, and the NOE data all yield a consistent picture. The pH 3.8 native state is well-folded and adopts a structure similar to

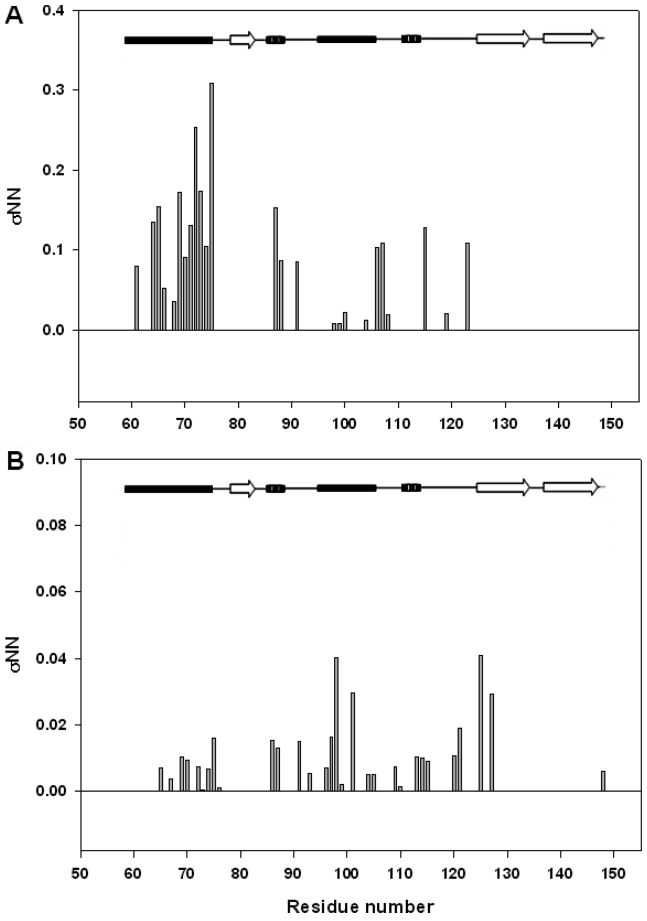

FIGURE 9: Summary of sequential d_{NN}(i,i + 1) NOEs observed for CTL9 at pH 3.8. NOE peak volumes are normalized as the ratio of the volumes of the observed d_{NN}(i,i + 1) NOE crosspeaks over those of the diagonal peaks. (A) The pH 3.8 native state NOEs, and (B) the pH 3.8 unfolded state NOEs. Regions where overlapping or ambiguous NOEs were observed or where no NOEs were observed are left blank. A schematic representation of the secondary structural elements in the native state is displayed on the top.

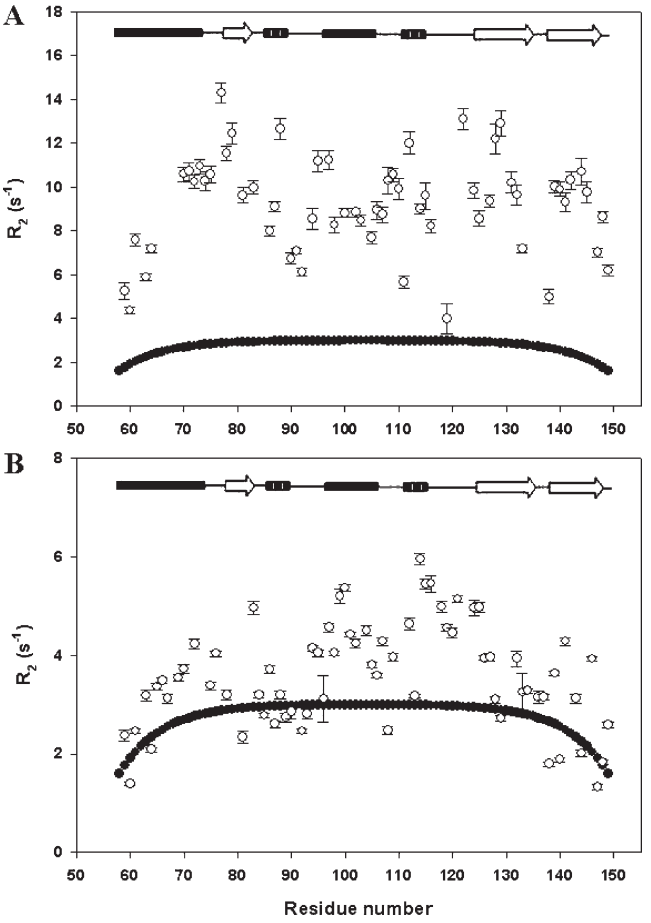

FIGURE 10: Plots of ^{15}N R_2 rates for the pH 3.8 native state (A), the pH 3.8 unfolded state (B) of CTL9. The solid line is the best fit to the phenomenological model of Schwalbe and co-workers (*3*).

the native state at neutral pH. The pH 3.8 unfolded state contains significant residual native and non-native structure. The two regions which are α-helical in the native state have a clear tendency to populate helical φ and ψ angles in the Ramachandran plot. The region that connects the second α-helix and the second β-strand adopts an irregular loop conformation in the native state, but displays a significant propensity to form non-native helical structure in the pH 3.8 unfolded state.

It is natural to inquire if significant unfolded state structure is to be routinely expected in the absence of denaturant. The question is difficult to answer since a limited number of proteins have had their unfolded states examined under near native conditions. Nonetheless, an examination of the literature suggests that deviations from random coil behavior are common and that unfolded states are often compact and can have a noticeable propensity to adopt secondary structure and form hydrophobic clusters (*10*, *12*). There are even examples of energetically significant electrostatic interactions in unfolded states (*4*, *5*, *13*, *16*, *47*−*55*). A second question concerns the presence of non-native structure. Again there are relatively few studies which directly address this point; however, non-native helical structure has been found in the drkN SH3 domain, in the ACBP domain, and as described here in CTL9 (*15*, *56*). Non-native helical structure may be more common than non-native β-sheet structure since helical geometry involves more local interactions which could be stabilized by transient longer range interactions present in a compact unfolded state (*57*−*59*).

The propensity to populate secondary structure in the pH 3.8 unfolded state is noticeably higher than that observed at pH 2.0 (*14*). This correlates with hydrodynamic measurements which have shown that the pH 2.0 unfolded state is more expanded than the pH 3.8 unfolded state (*13*). Thus, as one expects, the tendency to adopt native-like structure in the unfolded state decreases as solution conditions are changed to disfavor the native state. In principle, it is possible that the decreased propensity to populate helical structure in the pH 2.0 vs pH 3.8 unfolded state is due to a reduced intrinsic helical propensity at low pH. However, calculations performed using AGADIR (*60*), an algorithm that calculates intrinsic helical propensity without consideration of the potential effects of tertiary interactions, show that this is not the case. The calculations show only a weak intrinsic helicity for the relevant CTL9 regions (∼2−5%) that decreases only slightly (to ∼1−2%) upon moving from pH 3.8 to 2.0. This very small pH-dependent decrease cannot account for the sizable differences observed here between the residual helix populations of the pH 2.0 and the pH 3.8 unfolded states. Instead, given the

somewhat compact nature of both unfolded states, it may be that transient long-range interactions stabilize helical structure in both states, with the greater compactness of the pH 3.8 state affording a greater degree of helix stabilization. Precedent for such an effect is provided in studies of the acid unfolded state of apomyoglobin (*59*, *61*, *62*).

Residual structure in the unfolded state can reduce the conformational search in the early stages of protein folding, but it is important to realize that residual structure can speed up folding, have no effect on the folding rate, or even slow folding. If residual unfolded structure is no more developed in the transition state for folding than it is in the unfolded state, it will have no effect on the folding rate. This is true whether or not the structure represents native or non-native interactions (*55*, *63*, *64*). Unfolded state structure will slow folding if it must be disrupted prior to cross the transition state barrier.

Unfolded state structure can also affect the mutational analysis of the transition state for folding. This approach, traditionally denoted as Φ-value analysis, involves accessing the effect of a mutation on the overall stability, $\Delta\Delta G^0$, and on the activation free energy for folding, $\Delta\Delta G^{\ddagger}$. The dimensionless ratio $\Delta\Delta G^{\ddagger}/\Delta\Delta G^0$ gives the Φ-value. If the mutation does not alter the free energy of unfolded ensemble, then the ratio is just equal to the change in the free energy of the transition state divided by the change in the free energy of the native state. In this case, Φ-values have a simple quasi structural interpretation. $\Phi = 1$ represents an interaction which is as well developed in the transition state as in the native state, while $\Phi = 0$ represents an interaction that is no more developed in the transition state than it is in the unfolded state. Φ-values can obviously still be measured when mutations perturb the energetics of the unfolded state, but their simple structural interpretation is more ambiguous. For example, small Φ-values are traditionally assumed to represent lack of structure formation in the transition state; however, this is not always correct. If a particular interaction is well formed both in the transition state and in the unfolded state a small Φ-value can arise since Φ-values, in their simplest interpretation, report on the development of interactions in the transition state relative to the unfolded state. In such cases, a mutation may perturb the free energy of both the transition and unfolded states, giving rise to a Φ-value that is zero or small. Destabilization of the unfolded state via such a mutant will decrease the Φ-value relative to the value which would have been measured if the mutation had not affected unfolded state energetics (*23*). It is exceptionally difficult to experimentally detect when unfolded state effects may impact Φ-values but the observation of significant native and non-native helical propensities in the unfolded state of CTL9 indicates that such effects might arise in the case of this protein. If so, our view of how CTL9 folds may need to be reassessed carefully. Along these lines, it is interesting to note that a previously reported Φ-value analysis of the transition state of CTL9 demonstrated that the largest Φ-values are found in the β-hairpin region comprised of β-strand 2, β-strand 3, and the loop that connects the two strands, while the helical regions displayed small Φ-values.

ACKNOWLEDGMENT

We thank Prof. David Hoffman for providing the pH 5.5 native state assignments of CTL9, and we thank Dr. Shibani Bhattacharya and Mr. Wenli Meng for help with the NMR experiments. We also acknowledge Mr. Jian Feng for assistance with the calculation of the intrinsic relaxation rates.

SUPPORTING INFORMATION AVAILABLE

A figure showing the results of the AGADIR analysis and a table of the assignments for the pH 3.8 native state and the pH 3.8 unfolded state. This material is available free of charge via the Internet at http://pubs.acs.org

REFERENCES

1. Baldwin, R. L. (2002) A new perspective on unfolded proteins. *Adv. Protein Chem. 62*, 361–367.
2. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197–208.
3. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range interactions within a nonnative protein. *Science 295*, 1719–1722.
4. Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins. *J. Mol. Biol. 353*, 174–185.
5. Pace, C. N., Alston, R. W., and Shaw, K. L. (2000) Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci. 9*, 1395–1398.
6. Dill, K. A., and Shortle, D. (1991) Denatured states of proteins. *Annu. Rev. Biochem. 60*, 795–825.
7. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, D. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry 40*, 3561–3571.
8. Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739–756.
9. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasen, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 12491–12496.
10. Dyson, J. H., and Wright, P. E. (2002) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv. Protein Chem. 62*, 311–340.
11. Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A. 101*, 12497–12502.
12. Mittag, T., and Forman-Kay, J. D. (2007) Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol. 17*, 3–14.
13. Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterizatoin of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 349*, 839–846.
14. Shan, B., Bhattacharya, S., Eliezer, D., and Raleigh, D. P. (2008) The low-pH unfolded state of the C-terminal Domain of the ribosomal protein L9 contains significant secondary structure in the absence ofdenaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry 47*, 9565–9573.
15. Marsh, J. A., Neale, C., Jack, F. E., Choy, W. Y., Lee, A. Y., Crowhurst, K. A., and Forman-Kay, J. D. (2007) Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J. Mol. Biol. 367*, 1494–1510.
16. Cho, J.-H., Sato, S., Horng, J.-C., Anil, B., and Raleigh, D. P. (2008) Electrostatic interactions in the denatured state ensemble: Their effect upon protein folding and protein stability. *Arch. Biochem. Biophys. 469*, 20–28.
17. Gillespie, J. R., and Shortle, D. (1997) Charaterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol. 268*, 170–184.

18. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc. 123*, 11341–11352.
19. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1995) Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry 34*, 868–878.
20. Yang, D., Mok, Y.-K., Muhandiram, D. R., Forman-Kay, J. D., and Kay, L. E. (1999) ^{1}H-^{13}C dipole-dipole cross-correlated spin relaxation as a probe of dynamics in unfolded proteins: Application to the DrkN SH3 domain. *J. Am. Chem. Soc. 121*, 3555–3556.
21. Crowhurst, K. A., Choy, W.-Y., Mok, Y.-K., and Forman-Kay, J. D. (2003) Corrigendum to the paper by Mok et al. (1999) NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol. 329*, 185–187.
22. Anil, B., Craig-Schapiro, R., and Raleigh, D. P. (2006) Design of a hyperstable protein by rational consideration of unfolded state interactions. *J. Am. Chem. Soc. 128*, 3144–3145.
23. Cho, J.-H., and Raleigh, D. P. (2006) Denatured state effects and the origin of nonclassical values in protein folding. *J. Am. Chem. Soc. 128*, 16492–16493.
24. Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2006) Mutational analysis of the folding transition state of the C-terminal domain of ribosomal protein L9: A protein with an unusual β-sheet topology. *Biochemistry 46*, 1013–1021.
25. Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual alpha-beta topology: the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 345*, 163–173.
26. Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterizatoin of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 349*, 839–846.
27. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson. A. 102*, 241–245.
28. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson. 96*, 432–440.
29. Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H2O saturation. *J. Magn. Reson. A 109*, 129–133.
30. Clubb, R. T., Thanabal, V., and Wagner, T. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue ^{1}HN, ^{15}N, and $^{13}C'$ chemical shifts in ^{15}N—^{13}C-labelled proteins. *J. Magn. Reson. 97*, 213–217.
31. Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B. 101*, 201–205.
32. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. B. 103*, 203–216.
33. Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly ^{13}C/^{15}N-enriched proteins. *J. Biomol. NMR. 3*, 185–204.
34. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR. 4*, 845–858.
35. Grzesiek, S., Wingfield, P., Stahl, S., Kaufman, J. D., and Bax, A. (1995) Four-dimensional ^{15}N-separated NOESY of slowly tumbling perdeuterated ^{15}N-enriched proteins. Application to HIV-1 Nef. *J. Am. Chem. Soc. 117*, 9594–9595.
36. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*, 5984–6003.
37. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe-a multi-dimensional spectral processing system based on Unix Pipes. *J. Biomol. NMR. 6*, 277–293.
38. Johnson, B. A. (2004) Using NMRView to visualize and analyze NMR spectra of macromolecules. *Methods Mol. Biol. 278*, 313–352.
39. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc. 123*, 2970–2978.
40. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biol. NMR 5*, 67–81.
41. Spera, S., and Bax, A. (1991) Empirical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc. 113*, 5490–5492.
42. Wishart, D., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol. 22*, 311–333.
43. Eliezer, D. (2007) Characterization residual structure in disordered protein states using nuclear magnetic resonance. *Methods Mol. Biol. 350*, 49–67.
44. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Wright, P. E., and Dyson, H. J. (2001) Random coil chemical shifts in acidic 8M urea: Implementation of random coil shift data in NMRView. *J. Biol. NMR 18*, 43–48.
45. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci. 15*, 2795–2804.
46. Schwalbe, H. F., K., M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry 36*, 8977–8991.
47. Bowler, B. E. (2007) Thermodynamics of protein denatured states. *Mol. BioSyst. 3*, 88–89.
48. Whitten, S. T., and Garcia-Moreno, B. (2000) pH dependence of stability of staphylococcal nuclease: Evidence of substantial electrostatic interactions in the denatured state. *Biochemistry 39*, 14292–14304.
49. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR. *Chem. Rev. 104*, 3607–3622.
50. Religa, T. L., Markson, J. S., Mayor, U., Freund, S. M. V., and Fersht, A. R. (2005) Solution structure of a protein denatured state and folding intermediate. *Nature (London) 437*, 1053–1056.
51. Guzman-Casado, M., Parody-Morreale, A., Robic, S., Marqusee, S., and Sanchez-Ruiz, J. M. (2003) Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of thermus thermophilus ribonuclease H. *J. Mol. Biol. 329*, 731–743.
52. Kortemme, T., Kelly, M. J. S., Kay, L. E., Forman-Kay, J. D., and Serrano, L. (2000) Similarities between the spectrin SH3 domain denatured state and its folding transition state. *J. Mol. Biol. 297*, 1217–1229.
53. Tan, Y. J., Oliveberg, M., Davis, B., and Fersht, A. R. (1995) Perturbed pKa-values in the denatured states of proteins. *J. Mol. Biol. 254*, 980–992.
54. Swintkruse, L., and Robertson, A. D. (1995) Hydrogen-bonds and the pH-dependence of ovomucoid 3rd domain stability. *Biochemistry 34*, 4724–4732.
55. Cho, J. H., and Raleigh, D. P. (2006) Electrostatic interactions in the denatured state and in the transition state for protein folding: Effects of denatured state interactions on the analysis of transition state structure. *J. Mol. Biol. 359*, 1437–1446.
56. Kristjansdottir, S., Lindorff-Larsen, K., Fieber, W., Dobson, C. M., Vendruscolo, M., and Poulsen, F. M. (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *J. Mol. Biol. 347*, 1053–1062.
57. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature Struc.Biol. 5*, 148–155.
58. Lietzow, M. A., Jamin, M., Dyson, H. J., and Wright, P. E. (2002) Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol. 322*, 655–662.
59. Felitsky, D. J., Lietzow, M. A., Dyson, H. J., and Wright, P. E. (2008) Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. *Proc. Natl. Acad. Sci. USA. 105*, 6278–6283.
60. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) Elucidating the folding problem of α-helices: Local motifs, long-range electrostatics, ionic strength dependence and prediction of NMR parameters. *J. Mol. Biol. 284*, 173–191.
61. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol. 5*, 148–155.

62. Lietzow, M. A., Jamin, M., Dyson, H. J., and Wright, P. E. (2002) Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol. 322*, 655–662.
63. Luisi, D. L., Kuhlman, B., Sideras, K., Evans, P. A., and Raleigh, D. P. (1999) Effects of varying the local propensity to form secondary structure on the stability and folding kinetics of a rapid folding mixed α/β protein: characterization of a truncation mutant of the N-terminal domain of the ribosomal protein L9. *J. Mol. Biol. 289*, 167–174.
64. Daggett, V., Li, A. J., and Fersht, A. R. (1998) Combined molecular dynamics and Phi-value analysis of structure-reactivity relationships in the transition state and unfolding pathway of barnase: Structural basis of Hammond and anti-Hammond effects. *J. Am. Chem. Soc. 120*, 12740–12754.