Stony Brook University

NMR studies of the unfolded states of proteins under native and strongly denaturing conditions

A Dissertation Presented

by

Bing Shan

to

The Graduate School
in Partial Fulfillment of Requirements
for the Degree of
Doctor of Philosophy
in

Chemistry

Stony Brook University
December 2009

Stony Brook University

Bing Shan

We, the dissertation committee for the above candidate for the **Doctor of Philosophy** degree, hereby recommend acceptance of this dissertation.

Daniel P. Raleigh, Ph. D., Advisor
Professor of Chemistry, Department of Chemistry, Stony Brook University

Peter J. Tonge, Ph. D., Chairperson
Professor of Chemistry, Department of Chemistry, Stony Brook University

Jin Wang, Ph. D., Third Member
Professor of Chemistry, Department of Chemistry, Stony Brook University

David Eliezer, Ph.D., Outside Member
Associate Professor of Biochemistry, Weill Medical College of Cornell University

This dissertation is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of Dissertation

NMR studies of the unfolded states of proteins under native and strongly denaturing conditions

by

Bing Shan

Doctor of Philosophy
in
Chemistry

Stony Brook University
2009

There is considerable interest in the properties of the unfolded states of proteins, particularly unfolded states that are populated under native conditions. Interest in the unfolded state ensemble reflects the fact that it is the starting point for protein folding as well as the reference state for protein stability studies and can be the starting state for pathological protein aggregation. Unfolded states can be populated under a wide range of conditions but the unfolded state which is relevant for folding is the ensemble populated in the absence of denaturant under native conditions. In general this state cannot be directly probed because the free energy balance of folding strongly favors the folded state. The C-terminal domain of the ribosomal protein L9 (CTL9) undergoes a pH-induced unfolding transition and the native and the unfolded states are equally populated at pH 3.8 in the absence of denaturant. The two states are in slow exchange on the NMR time scale, allowing them to be studied under identical conditions. Detailed characterization of the pH 3.8 unfolded state reveals a compact unfolded state with both native and non-native structure. In contrast, the pH 2.0 unfolded state and the urea unfolded state have much less residual structure. The implications of unfolded state structure on the analysis of protein folding transition states are discussed.

Protein cold denaturation has attracted recent attention because of the insight it may provide into the origins of the cooperativity of protein folding and the nature of partially folded states. Unfortunately, studies of protein cold denaturation have been hindered because the cold denatured state is normally difficult to experimentally access. The cold denatured state of CTL9 can be populated under native like conditions by taking advantage of a destabilizing point mutation which leads to cold denaturation at temperatures above zero ^{o}C. The cooperativity of cold denaturation is tested and the structural propensities of the cold denatured state were analyzed by NMR spectroscopy. The cold denatured state of CTL9 contains significant native and non-native secondary structure, but

also contains transiently formed long range interactions. The implications for protein folding and for studies of cold denatured state are discussed.

Table of Contents

List of figures ix
List of tables xi
List of symbols and abbreviations xii
List of publications xiv

1.1. Protein folding 1
1.1.1. Mechanisms of protein folding 1
1.1.2. Cooperativity of protein folding 3
1.1.3. Thermodynamics of protein folding 3
1.1.4. Kinetics of protein folding 4
1.2. The unfolded states of proteins 5
1.2.1. The significance of the unfolded states of proteins 5
1.2.2. Characteristics of natively unfolded proteins proteins 6
1.2.3. Experimental methods to address unfolded proteins 7
1.3. Protein cold denaturation 7
1.4. The ribosomal protein L9 8
1.4.1. The structure and function of the ribosomal protein L9 8
1.4.2. The C-terminal domain of the ribosomal protein L9 9
1.5. The aims of this dissertation 10

2. The low pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low pH urea unfolded state 16
2.1. Introduction 17
2.2. Materials and methods 18
2.2.1. Mutagenesis, protein expression and purification 18
2.2.2. NMR sample preparations 18
2.2.3. NMR spectroscopy 18
2.2.4. Measurement of transverse relaxation rates 19
2.2.5. Preparation of the spin-labeled sample 19
2.2.6. Data processing and analysis 20
2.3. Results 21
2.3.1. Hydrodynamic properties of the unfolded states of CTL9 21
2.3.2. Sequence specific assignments of CTL9 in the pH 2.0 acid unfolded state and in the pH 2.5, 7.6 M urea unfolded state 21
2.3.3. Analysis of ^{13}C, ^{1}H and ^{15}N chemical shifts provide evidence for native secondary structure in the unfolded state in the absence of denaturant 22
2.3.4. ^{15}N R_2 relaxation analysis of the low pH unfolded state and the low pH urea unfolded state of CTL9 25
2.3.5. Mapping long range interactions in the unfolded states of CTL9 25
2.4 Discussion and conclusions 26

3. The unfolded state of the C-terminal domain of the ribosomal protein L9 contains both native and non-native structure 44
3.1. Imtroduction 44
3.2. Materials and methods 46
3.2.1. Protein expression and purification 46
3.2.2. NMR experiments 46
3.3. Results 47
3.3.1. Sequence specific assignments of the native state and the unfolded state of CTL9 at pH 3.8 47
3.3.2. Analysis of the ^{13}C and ^{1}H chemical shifts indicates significant native and non-native secondary structure in the unfolded state 48
3.3.3. Analysis of chemical shift differences and secondary structural propensity scores indicates significant helical structure in the unfolded state 50
3.3.4. Amide proton NOEs confirm the presence of helical structure in the unfolded state. 51
3.3.5. ^{15}N R_2 relaxation analysis of the pH 3.8 native state and the pH 3.8 unfolded state 52
3.4. Discussion 52

4. The cold denaturation of the C-terminal domain of the ribosomal protein L9 is a cooperative, two-state process. 67
4.1. Introduction 67
4.2. Materials and methods 67
4.2.1. Mutagenesis, protein expression and purification 67
4.2.2. Circular dichroism spectroscopy 68
4.2.3. Nuclear magnetic resonance experiments. 68
4.2.4. Stopped-flow fluorescence 68
4.3. Results 69
4.3.1. Thermodynamics of cold denaturation of I98A CTL9 69
4.3.2. Temperature dependent NMR experiments suggest two macroscopic states associated with the cold denaturation of I98A CTL9 70
4.3.3. The cold denaturation of I98A CTL9 is cooperative 70
4.4. Discussion 71

5. The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure 82
5.1. Introduction 83
5.2. Materials and methods 84
5.2.1. Protein expression and purification 84
5.2.2. Circular dichroism spectroscopy. 84
5.2.3. Stopped-flow fluorescence 84
5.2.4. NMR experiments. 84

5.3. Results 85
5.3.1. The I98A mutant of CTL9 adopts the same structure as wild-type CTL9 but undergoes cold denaturation above 0 ^{o}C 85
5.3.2. Chemical shift analysis provides evidence of native and non-native secondary structure in the cold denatured state 86
5.3.3. Residual dipolar couplings confirm α-helical secondary structure in the cold denatured state Amide proton NOEs provide additional evidence of native like structure in the cold denatured state 87
5.3.4. Amide proton NOEs are consistent with native like structure in the cold denatured state 88
5.3.3. ^{15}N transverse relaxation rate analysis of the cold denatured state 89
5.4. Discussion and conclusions 89

6. Correspondence between the overall dimensions and the amount of the secondary structure in the unfolded states of proteins. 104

References 109

Appendix 1. CTL9 backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments at pH 2.0 123
Appendix 2. CTL9 backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments at pH 2.5, in 7.6 M urea. 125
Appendix 3. ^{15}N R_2 relaxation rates for the pH 2.0 and pH 2.5, 7.6 M urea unfolded states of CTL9 and calculated R_2 rates using random coil model 127
Appendix 4. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-^{1}H crosspeaks in the HSQC spectrum of K74C, K109C and D119C mutants of CTL9 at pH 2.0 129
Appendix 5. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-^{1}H crosspeaks in the HSQC spectrum of K74C, K109C and D119C mutants of CTL9 at pH 2.0 in 8M urea 131
Appendix 6. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of the pH 3.8 native state of CTL9 133
Appendix 7. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of the pH 3.8 unfolded state of CTL9 135
Appendix 8. ^{15}N R_2 relaxation rates for the pH 3.8 native state and the pH 3.8 unfolded states of CTL9 and calculated R_2 rates using random coil model 137
Appendix 9. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of I98A CTL9 at pH 5.7 and 12 ^{o}C 139
Appendix 10. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of wild type CTL9 at pH 5.7 and 12 ^{o}C 141
Appendix 11. ^{15}N R_2 relaxation rates for I98A CTL9, wt-CTL9 at pH 5.7, 12 ^{o}C and calculated R_2 rates using random coil model. 143

Appendix 12. ^{1}H-^{15}N RDCs for I98A CTL9 and wt-CTL9 at pH 5.7, 12 ^{o}C 145

List of Figures

Figure 1-1: Three dimensional representation of a folding funnel. 11
Figure 1-2: An energy diagram representing a two-state protein folding reaction. 12
Figure 1-3: Plot of the free energy change of the protein unfolding reaction versus temperature 13
Figure 1-4: A ribbon diagram of the ribosomal protein L9 14
Figure 1-5: Ribbon diagrams of the C-terminal domain of protein L9 15
Figure 2-1: A ribbon diagram and the primary sequence of CTL9. 30
Figure 2-2: Reaction scheme of MTSL with cysteine sidechain and TCEP 31
Figure 2-3: The ^{15}N-^{1}H HSQC spectrum of the low pH unfolded state and the low pH urea unfolded state of CTL9. 32
Figure 2-4: Plots of $^{1}H_{\alpha}$ secondary shifts of the pH 3.8 native state, the low pH unfolded state and the low pH urea unfolded state of CTL9.. 33
Figure 2-5: Plots of ^{13}C secondary shifts of the low pH unfolded state of CTL9 34
Figure 2-6: Plots of ^{13}C secondary shifts of the low pH urea unfolded state of CTL9 35
Figure 2-7: Plots of the $^{13}C_{\alpha}$-$^{13}C_{\beta}$ secondary shifts of the low pH unfolded state and the low pH urea unfolded state of CTL9. 36
Figure 2-8: SSP analysis of the low pH unfolded state and the low pH urea unfolded state of CTL9 calculated using $^{13}C_{\alpha,}$ $^{13}C_{\beta}$ and $^{1}H_{\alpha}$ chemical shifts 37
Figure 2-9: SSP analysis of the low pH unfolded state and the low pH urea unfolded state of CTL9 calculated using $^{13}C_{\alpha}$, $^{13}C_{\beta}$ chemical shifts 38
Figure 2-10: ^{15}N R_2 rates for the low pH unfolded state and the low pH urea unfolded state of CTL9 39
Figure 2-11: The ^{15}N-^{1}H HSQC spectra of wild-type CTL9, K74C CTL9, K109C CTL9 and D119C CTL9 40
Figure 2-12: Plot of paramagnetic relaxation enhancement for the low pH, 8M urea unfolded state of CTL9 41
Figure 2-13: Plot of paramagnetic relaxation enhancement for the low pH unfolded state of CTL9. 42
Figure 2-14: AGADIR analysis of the primary sequence of CTL9 at pH 2.0, 298 ^{o}K. 43
Figure 3-1: A ribbon diagram and the primary sequence of CTL9 56
Figure 3-2: ^{15}N-HSQC spectrum of CTL9 recorded at pH 3.8 and 25 ^{o}C 57
Figure 3-3: Plots of $^{1}H\alpha$ secondary shifts of the pH 3.8 native state and the pH 3.8 unfolded state of CTL9 58
Figure 3-4: Plots of ^{13}C secondary shifts for the pH 3.8 native state of CTL9 59
Figure 3-5: Plots of ^{13}C secondary shifts for the pH 3.8 unfolded state of CTL9 60
Figure 3-6: Plots of the $^{13}C_{\alpha}$-$^{13}C_{\beta}$ secondary shifts of the pH 3.8 native state and the pH 3.8 unfolded state of CTL9 61
Figure 3-7: SSP analysis of the pH 3.8 native state and the pH 3.8 unfolded state of CTL9 62
Figure 3-8: Correlation plots of the SSP scores of the pH 3.8 native state, the pH 3.8 unfolded state and the low pH urea unfolded state of CTL9. 63

Figure 3-9: Summary of sequential amide proton NOEs observed in the pH 3.8 native state and the pH 3.8 unfolded state of CTL9. 64
Figure 3-10: Plots of ^{15}N R_2 rates for the pH 3.8 native state and the pH 3.8 unfolded state of CTL9. 65
Figure 3-11: AGADIR analysis of the primary sequence of CTL9 at pH 2.0 with 20 mM salt and at pH 3.8 with 120 mM salt. 66
Figure 4-1: A ribbon diagrams of CTL9. 75
Figure 4-2: CD monitored pH-induced denaturation of I98A CTL9..... 76
Figure 4-3: CD monitored thermal denaturation experiments of I98A CTL9 as a function of pH 77
Figure 4-4: ^{1}H-^{15}N HSQC spectra of I98A CTL9 as a function of temperature..... 78
Figure 4-5: Chevron plots of the I98A CTL9 at pD 6.6 79
Figure 4-6: Far-UV CD spectra of I98A mutant of CTL9 as a function of temperature. 80
Figure 4-7: SVD analysis of the temperature dependent CD spectra of I98A CTL9 81
Figure 5-1: Ribbon diagrams of CTL9 and the CD and NMR monitored thermal denaturation of I98A CTL9. 94
Figure 5-2: ^{15}N-HSQC spectrum of I98A CTL9 recorded at pH 5.7, 12 ^{o}C..... 95
Figure 5-3: The fluorescence change as a function of time in the pH jump stopped flow experiments of I98A CTL9 96
Figure 5-4: Plots of $^{13}C_{\alpha}$, $^{13}C_{\beta}$, ^{13}CO and $^{1}H_{\alpha}$ secondary shifts for the cold denatured state of I98A CTL9 97
Figure 5-5: Plots of the $^{13}C_{\alpha}$-$^{13}C_{\beta}$ secondary shifts of wild type CTL9 and the cold denatured state of I98A CTL9 at pH 5.7, 12 oC 98
Figure 5-6: SSP analysis of wild type CTL9 and I98A CTL9 at pH 5.7, 12 ^{o}C 99
Figure 5-7: ^{1}H-^{15}N residual dipolar couplings for wild type CTL9 and I98A CTL9 100
Figure 5-8: Summary of NOEs observed for I98A CTL9 at pH 5.7, 12 ^{o}C 101
Figure 5-9: Plots of 15N R2 relaxation rates for wild type CTL9 and I98A CTL9 102
Figure 5-10: Correlation plots of the SSP scores of the cold denatured state of I98A CTL9 and those of the pH 3.8 unfolded state of wild type CTL9. 103
Figure 6-1: Correlation between compactness and the amounts of secondary structure determined by CD experiments..... 107
Figure 6-2: Correlation between the degree of compactness and the relative amounts of secondary structure calculated using SSP scores 108

List of Tables

Table 2-1: Hydrodynamic radii for the native, acid unfolded and urea unfolded states of CTL9 29
Table 3-1: Average secondary shifts and SSP values in secondary structural elements of CTL9 55
Table 4-1: Thermodynamic parameters for the unfolding of wild-type CTL9 and the I98A mutant determined from CD monitored thermal denaturation experiments 72
Table 4-2: Estimated mid points of thermal and cold denaturation of I98A CTL9 at different pH values 73
Table 4-3: Kinetic parameters for the unfolding and refolding of I98A CTL9 at various temperatures determined by stopped-flow experiments 74
Table 5-1: Average secondary shifts and SSP values in secondary structural elements of I98A CTL9 at pH 5.7, 12 ºC 92
Table 5-2: Average secondary shifts and SSP values in secondary structural elements of wild type CTL9 at pH 5.7, 12 ºC 93
Table 6-1: Average SSP scores, CD ellipticity at 222 nm and radii of hydration for various conformational states of CTL9 106

List of Symbols and Abbreviations

CD	circular dichroism
CTL9	The isolated C-terminal domain of the ribosomal protein L9 from *Bacillus stearothermophilus*
DSS	4, 4-dimethyl-4-silapentane-1-sulfonic acid
F	The folded state
ΔC°_{p}	Change in heat capacity upon unfolding
$\Delta G^{\ddagger\circ}$	Activation free energy for folding
ΔG°	Equilibrium free energy of folding
ΔH°	Standard state enthalpy of unfolding
ΔS°	Standard state entrolpy of unfolding
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum correlation
IPAP-HSQC	Inphase and antiphase-HSQC
IPTG	Isopropylthio-β-D-galactoside
k_{obs}	Rate constant for exchange
k_f	Rate constant for folding
k_u	Rate constant for unfolding
m	Slope of ΔG° vs. denaturant concentration
m_f	Slope of ln k_f vs. denaturant concentration
m_u	Slope of ln k_u vs. denaturant concentration
NOESY	Nuclear Overhauser effect spectroscopy
NMR	Nuclear magnetic resonance
MALDI-TOF	Matrix assisted laser desorption ionization-time of flight
MTSL	*S*-(2, 2, 5, 5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
PFG-NMR	Pulsed field gradient nuclear magnetic resonance
R	Universal gas constant
R_g	Radius of gyration
R_h	Radius of hydration
RDC	Residual dipolar couplings
SAXS	Small angle X-ray scattering
SSP	Secondary structural propensity
TCEP	Tris (2-carboxylethyl) phosphine
TFA	Trifluoroacetic acid
T	Temperature
T_c	The temperature at the midpoint of the cold denaturation
T_m	The temperature at the midpoint of the thermal unfolding transition
TOCSY	Total correlation spectroscopy
U	The unfolded state
[den]	Concentration of denaturant
[θ]	Mean residue ellipticity

Acknowledgements

I would like to thank my adviser, Professor Daniel. P. Raleigh for introducing me to the field of protein folding and NMR spectroscopy. I appreciate all his scientific guidance, encouragement, and full support throughout the course of my research at Stony Brook.

I would like to thank Professor Peter J. Tonge, Professor Jin Wang, and Professor David Eliezer for their contributions to my dissertation. I am grateful to have an opportunity to collaborate with Professor David Eliezer. I have great experience in working with Dr. Shibani Bhattacharya at the New York Structural Biology Center who has made great contributions to my NMR experiments. I also thank Prof. Scott Brewer for collaboration in the cold denaturation project. I wish to thank the chemistry department of Stony Brook University for providing me with the opportunities for pursuing my advanced education.

I am grateful for the friendship with the past and present members of the Raleigh group: Dr. Ying Li, Dr. Yuefeng Tang, Dr. Yuan Bi, Dr. Sylvia Tracz, Dr. Benben Song, Dr. Andisheh Abedini, Dr. Jae-hyun Cho, Dr. Burcu Anil, Dr. Ruchi Gupta, Konstantine Aprilakis, Humeyra Taskent, Fanling Meng, Wenli Meng, Shifeng Xiao, Peter Marek, Vadim Patsalo, Ping Cao, Ivan Peran, Trisha Barua, Bowu Luan, Hui Wang, Ling-Hsien Tu. I specially thank Dr. Ying Li for introducing me to various experimental techniques and I would like to thank Mr. Wenli Meng, Mr. Shifeng Xiao and Mr.Vadim Patsalo for several productive collaborations.

Last but not the least I would like to thank my family for their love, support and understanding.

List of Publications

1. Li, Y., **Shan, B.** & Raleigh, D. P. The Cold Denatured State Is Compact but Expands at Low Temperatures: Hydrodynamic Properties of the Cold Denatured State of the C-terminal Domain of L9. *Journal of Molecular Biology*. 368, 256-262, 2007.

2. **Shan, B.**, Bhattacharya, S., Eliezer, D. & Raleigh, D. P. The Low pH Unfolded State of the C-terminal Domain of the Ribosomal Protein L9 Contains Significant Secondary Structure in the Absence of Denaturant but Is No More Compact than the Low pH Urea Unfolded State. *Biochemistry*. 47, 9565-9573, 2008.

3. Bi, Y., Cho, J., Kim, E., **Shan, B.**, Schindelin, H. & Raleigh, D. P. Rational Design, Structural and Thermodynamic Characterization of a Hyperstable Variant of the Villin Headpiece Subdomain. *Biochemistry*. 46, 7497-7505, 2007.

4. **Shan, B.**, Eliezer, D. & Raleigh, D. P. The Unfolded State of the C-terminal Domain of the Ribosomal Protein L9 Contains Both Native and Non-Native Structure. *Biochemistry*. 48, 4707-4719, 2009.

5. Xiao, S., Bi, Y., **Shan, B.** & Raleigh, D. P. Analysis of Core Packing in a Cooperatively Folded Miniature Protein: The Ultrafast Folding Villin Headpiece Helical Subdomain. *Biochemistry*. 48, 4607-4616, 2009.

6. Meng, W., **Shan, B.**, Tang, Y. & Raleigh, D. P. Native Like Structure in the Unfolded State of the Villin Headpiece Helical Subdomain, an Ultrafast Folding Protein. *Protein Science*. 18, 1692-1701, 2009.

7. **Shan**, B., McClendon, S., Rospigliosi, C., Eliezer, D. & Raleigh, D. P. The Cold Denatured State of the C-terminal Domain of Protein L9 Is Compact and Contains Both Native and Non-native Structure. *Submitted.*

1. Introduction

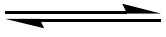
1.1 Protein folding

Proteins, the biochemical molecules which make up cells, organs and organisms, are translated from mRNA on the ribosome. Newly expressed proteins exist as unfolded polypeptides which lack three dimensional structure. The process through which unfolded polypeptide chains self-assemble into the native three dimensional structure is called protein folding. A major breakthrough in the study of protein folding occured in the 1950s, when Anfinsen demonstrated in his famous experiment (*1*) that the protein ribonuclease A (RNaseA) becomes a non-structured random polymer and loses it enzymatic activity when the disulfide bonds are reduced and the protein is denatured by urea. Subsequently, when the urea is removed and the –SH groups are oxidized to form disulfide bonds, RNaseA can fold back into its native structure spontaneously and regain its biological activity by more than 90%. The Anfinsen experiment revealed that all the information required to dictate the three dimensional structure is encoded in the primary amino acid sequence and that unfolded proteins can spontaneously refold but it did not provide any details about the mechanism of protein folding. A wide range of experiments have been performed on a large number of proteins, in an effort to understand the mechanism of protein folding, the physical basis of the stability of the folded state, and how to predict protein structure based on the amino acid sequence, but the folding process is still not understood in detail.

1.1.1 Mechanisms of protein folding

A fundamental problem of protein folding is how a given amino acid sequence folds into its native state spontaneously and if there is a defined pathway which the folding reaction follows. According to the famous "Levinthal paradox" (*2*), there exists a huge number of conformations in the unfolded state. A 100 residue protein will have 2^{100} possible conformations if only two conformations are allowed per residue. It would take far longer than the age of the universe for proteins to fold into their native structures by randomly sampling all of these conformations. However, proteins attain their unique native conformation on a time scale of typically seconds or even submilliseconds for some small proteins (*3-5*). So there must be a folding pathway or funnel shaped landscape which allows proteins to fold at high efficiency. Deciphering the mechanism of protein folding is of great importance because protein misfolding or incomplete folding is involved in many diseases including Alzheimer's disease, Parkinson's disease and type II diabetes. Understanding the folding pathway will also improve the production of correct folded recombinantly produced proteins for industrial and research applications.

Protein folding can proceed via a variety of different mechanisms and are often classified as either a two-state process or a three-state or multi-state process. The simplest case is the two-state process, i.e., a folding reaction which proceeds without the involvement of any detectable intermediates.



(Two-state mechanism)

In this case, the protein has only two thermodynamicly stable states, folded and unfolded, where U denotes the unfolded state of a protein and F refers to the native

folded state. Some proteins may form stable intermediate states which can be on or off the folding pathway. More complicated equations must be used to describe the equilibrium and kinetics of the three-state or multi-state process, especially if off pathway species are involved.

$$U \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons \cdots \rightleftharpoons F$$ (Three or Multi-state mechanism)

Several models have been developed to describe protein folding pathways. In reality, a protein folding process may incorporate the features of several models. The framework model (*6-9*) proposes that in the initial stages of protein folding, there is a certain degree of preformed residual secondary structure in the polypeptide chain which is independent of the tertiary structure. As folding proceeds, these elements of secondary structure diffuse and collide, whereby they hierarchically assembled into stabilized tertiary structure. The whole process occurs in a stepwise manner and thus greatly reduces the conformational search. The process is very similar to the diffusion and collision model which was originally developed to explain the folding of helical proteins (*10-11*). In fact, the framework model and the diffusion and collision model are often considered to be fundamentally equivalent. An alternative mechanism, the so called hydrophobic collapse model hypothesizes that the protein chain rapidly collapses to form a compact molten globule sate driven by hydrophobic interactions, which contains significant secondary structure but lacks tertiary structure. The compact molten globule state then rearranges to form the final stable native conformation (*12-16*). The framework model, the diffusion collision model and the hydrophobic collapse model all envisage a sequential folding pathway with a limited number of kinetic intermediates. On the other hand, the nucleation condensation model, proposed by Fersht (*17-18*), argues that the formation of the folding nucleus in the transition state by a cluster of key residues contributes to the initial rate-limiting step for folding. The nucleation sites may not be stable but they act as the template which facilitates the remaining the polypeptide chain to fold rapidly. The nucleation condensation model does not assume the presence of any kinetic intermediate and appear to be supported by experimental evidence from folding studies of small single domain proteins, such as chymotrypsin inhibitor-II (*19-21*) and barstar (*22*) which have been demonstrated to fold highly cooperatively.

More recently, a new model based on free energy landscapes, termed the folding funnel theory has been proposed (*23-27*). Landscape theory is a more general description of the protein folding pathway from the thermodynamic point of view, which imagines the free energy surface of a protein folding process as a free energy funnel, as shown in Figure 1-1. In the folding funnel, the vertical axis represents the free energy less the contribution from configuration entropy while the width of the funnel is related to the configuration entropy of the protein. The shape of the funnel is unique to a specific protein and is determined by the primary amino acid sequence. The unfolded states with a high degree of configuration entropy and relatively high free energy are located at the top of the funnel, which allows many distinct conformational states. The native state with a much smaller number of conformations but the lowest free energy resides at the bottom of the funnel. Instead of having smooth sides, the funnel may have rough surface with some local energy minimum, which generate kinetic intermediates on the folding pathway. It is important to note that folding free energy funnels are generally smooth especially when compared to processes such as glass transition. The significance of the funnel theory is that the protein folding may proceed via a variety of pathways, from the

unfolded state to the native state, instead of by a single fixed route. As plotted in Figure 1-1, which is adapted from reference (*26*), folding may follow the fast folding pathway (yellow), the slow folding pathway (red) or the slow folding pathway which returns to a less folded intermediate state and then adopts the fast folding pathway.

1.1.2 Cooperativity of protein folding

In the exploration of protein folding pathways, a great deal of effort has been devoted to the search for kinetic intermediates. However, Fersht et. al. founded in 1991 that chymotrypsin inhibitor-II (CI2) (*17, 19-20*), a small single-domain protein, folds in a two-state fashion and no stable kinetic intermediates with significant population could be detected. Subsequently, many small single-domain proteins have been demonstrated to fold in a highly cooperative process (*28-31*).

There are many experimental methods to test the cooperativity of folding. A classical approach is to monitor the folding or unfolding transition by two distinct spectroscopic properties, such as circular dichroism and fluorescence and plotting the normalized signals. If the folding is cooperative, the signals from the two different probes should overlap and generate identical thermodynamic parameters (*32*).

Kinetic experiments are also often used to study protein folding cooperativity. Most protein folding reactions are studied by stopped-flow method. The standard method for analyzing stopped-flow kinetic data is to plot the log of the observed rate (k_{obs}) vs the final concentration of denaturant after mixing, which is usually termed a Chevron plot. Protein stability is normally a linear function of denaturant concentration, thus a plot of $\ln(k_{obs})$ vs the concentration of denaturant is equivalent to a linear free energy relationship. Any deviation from linear behavior or roll-over at low concentration of denaturant in the Chevron plot, are indications of the presence of intermediates (*33-34*).

NMR spectroscopy, which provides structural and dynamic information at atomic level resolution, is particularly useful in delineating the structures of the intermediates. Chemical shifts, which are very sensitive to environmental changes, or the resonance intensities, are usually monitored as a function of denaturant concentration or temperature. Numerous studies have been reported, for example, Radford and coworkers have identified the conformational property of an amyloidogenic intermediate of human β_2-microglobulin (β_2m) using equilibrium unfolding experiments by ^{1}H -^{15}N NMR (*35*), while Wand et. al have utilized temperature dependent ^{1}H -^{13}C NMR to study the cooperativity of cold denaturation of encapsulated ubiquitin (*36*).

1.1.3 Thermodynamics of protein folding

Studying the thermodynamic properties of protein folding is crucial for understanding why the native state of a protein is stable. Protein stability is defined as the difference of the free energy between the unfolded state and the native state.

$$\Delta G^{o}_{N\text{-}D} = G^{o}_{D} - G^{o}_{N} \quad (1)$$

The most common methods to determine protein stability are chemical denaturation or thermal denaturation experiments. The unfolding free energy (ΔG^{o}) in the presence of chemical denaturant is usually linearly dependent on the concentration of the denaturant for chemical denaturation (*37-38*).

$$\Delta G^{o} = \Delta G^{o}(H_2O) - m\,[denaturant] \quad (2)$$

where ΔG^o (H_2O) is the unfolding free energy in the absence of denaturant. m is a parameter which is relevant to the change of solvent accessible surface area upon unfolding. For thermal denaturation, the unfolding free energy follows the Gibbs-Helmholtz equation:

$$\Delta G^o(T) = \Delta H^o(T_m)(1 - \frac{T}{T_m}) - \Delta C_p{}^o[(T_m - T) + T\ln(\frac{T}{T_m})] \quad (3)$$

where T_m is the midpoint of the thermal denaturation, defined as the temperature where there are equal amounts of folded and unfolded protein. ΔH^o (T_m) is the enthalpy change upon unfolding at T_m, and $\Delta C_p{}^o$ is the heat capacity change upon unfolding at T_m, which is also related to the change of solvent accessible surface area upon unfolding (*39*). Equation (3) requires that $\Delta C_p{}^o$ be independent of temperature which is a reasonable assumption over the temperature range normally studied.

Chemical denaturation and thermal denaturation experiments are usually monitored by circular dichroism (CD) or fluorescence spectroscopy. For a two-state folding system, the denaturation curves can be fit to obtain thermodynamic parameters. For chemical denaturation, the curves can be analyzed by a non-least square curve fitting equation:

$$\theta([den]) = \frac{(a_n + b_n[den]) + (a_d + b_d[den])\exp(-\Delta G_u{}^o([den])/RT)}{1 + \exp(-\Delta G_u{}^o([den])/RT)} \quad (4)$$

where a_n, b_n, a_d and b_d are parameters which defined the signals of the native state (N) and denatured state (D) at a given denaturant concentration. [den] denotes the concentration of denaturant and ΔG^o is the equilibrium free energy of unfolding. Thermal unfolding data can be fit to the following equation to obtain thermodynamic parameters:

$$\theta(T) = \frac{(a_n + b_n T) + (a_d + b_d T)\exp(-\Delta G_u{}^o(T)/RT)}{1 + \exp(-\Delta G_u{}^o(T)/RT)} \quad (5)$$

where a_n, b_n, a_d and b_d are parameters which defined the signals of the native state (N) and denatured state (D) at a given temperature. T_m is the heat induced unfolding midpoint temperature, $\Delta H^o(T_m)$ is the enthalpy change at T_m, and $\Delta C_p{}^o$ is the heat capacity change between the native and denatured states.

1.1.4 Kinetics of protein folding

Kinetic measurement of protein folding provides insights into the mechanism of folding. The folding events occur in a wide time frame ranging from microsecond to hundreds of seconds depending on the amino acid sequence and the size of the protein. The most common kinetic experiment is the stopped-flow experiment, which measures the observed rate of folding as a function of denaturant concentration. Similar to the equilibrium free energy, the folding rate and unfolding rate in a folding reaction also depend on the concentration of the denaturant, such as urea and guanidine hydrochloride:

$$\ln(k_{obs}) = \ln(k_{f(H_2O)}e^{m_f[den]/RT} + k_{u(H_2O)}e^{m_u[den]/RT}) \quad (6)$$

where [den] denotes the concentration of denaturant, $k_{f\,(H2O)}$ and $k_{u\,(H2O)}$ are the folding and unfolding rate in the absence of denaturant. m_f and m_u are constants which describe how $\ln(k_f)$ and $\ln(k_u)$ change as a function of the concentration of denaturant. m_f is related to the difference of the solvent accessible surface area between the unfolded state and the transition state, while m_u is related to the difference of the solvent accessible surface area

between the transition state and the native state. For a two-state folding reaction, $m_{eq} = m_u - m_f$, and the plot of k_{obs}([den]) versus denaturant concentration is a V-shaped curve, termed a Chevron plot. The folding and unfolding rate in the absence of denaturant can be obtained by fitting the Chevron plot using equation (6).

Traditional stopped-flow spectrometers typically have a dead-time on the order of a few milliseconds and thus are not applicable to fast folding proteins. For proteins which fold in less than a millisecond, fast mixing techniques such as continuous flow or fast initiation of folding process such as temperature jump experiments have to be employed. Continuous flow experiments can be used to detect kinetic process in the range of 50 microseconds to 1 millisecond (*40-41*). In temperature jump experiments (*42-43*), the equilibrium can be perturbed by a laser or a electrostatic discharge in nanoseconds. Thus temperature jump experiments are widely used to study very fast folding processes. Alternatively, dynamic NMR lineshape analysis can also be used to measure folding rates for fast folding proteins, provided that the exchange rates between the native state and the unfolded state are comparable to the NMR chemical shift time scale (*44-46*). NMR relaxation dispersion experiments provide a more general approach and extend the range of systems that can be probed by NMR (*47*).

1.2 The unfolded state of proteins

1.2.1 The significance of the unfolded state of proteins

Probing protein folding pathways requires the characterization of all species along the folding reaction, including the folded states, unfolded states, the transition state and any intermediates. Detailed knowledge of the native state has been available for decades. The transition states of protein folding are widely investigated using φ-value analysis (*48-49*). For a two-state folding reaction which proceeds via a single energy barrier as displayed in Figure 1-2, the φ-value is defined as the ratio of the change in the activation free energy caused by a mutation over the change in the equilibrium free energy due to that mutation.

$$\varphi = \frac{\Delta\Delta G^{\neq}}{\Delta\Delta G^{o}} \tag{7}$$

φ-value evaluates the development of a particular interaction in the transition state relative to the unfolded state. If a mutation destabilizes the native state as much as the transition state, the φ-value equals to 1, indicating that interactions involving in this particular residue is fully formed in the transition state. In the other extreme, if a mutation does not affect the folding rate yet destabilizes the native state, the φ-value equals to zero, indicating that the interactions being probed by the mutation is no more developed in the transition state than in the unfolded state. An essential prerequisite of φ-value analysis is that the mutation does not perturb the energetics of the unfolded state. When the energy of the unfolded state ensemble is changed by a mutation, the measured φ-value will deviate from the value where the unfolded state is not changed (*50*). As a consequence, the knowledge of the unfolded state is of significance on the understanding the transition state of folding.

Unfortunately relatively little is known about the unfolded state, especially unfolded states which are populated under physiological conditions (*51-53*). The interest in the unfolded state is largely fueled by the realization that the unfolded states are the proper reference state for protein folding. Knowledge of the unfolded state ensemble may

thus facilitate the understanding of protein stability and folding kinetics. The unfolded state ensemble can even be a target for protein engineering studies, since the net stability of the protein will be increased if the unfolded states ensemble can be destabilized (*54*). In addition, the unfolded state is the starting state for protein folding, and residual structure in the unfolded state is thought to reduce the initial conformational search and thus accelerate protein folding (*55*).

Protein aggregation and misfolding are now widely recognized as important and competing pathways to protein folding (*56*). Mechanistic studies suggest that protein aggregation and amyloid formation start from unfolded states or partially unfolded states (*57-58*). Residual structure in the unfolded state not only contributes to the initiation of folding, but may play a role in protein aggregation (*59-60*). Biophysical studies of amyloidogenic proteins such as α-synuclein (*61-62*), β_2-microglobulin (*63*), and lysozyme (*64*) revealed that there are partially formed elements of secondary structure or hydrophobic or aromatic clusters in the unfolded states of these proteins. Mutational studies (*60*) indicated that changes in the unfolded state structures may result in modulation of the kinetics of amyloid formation.

In the contrary to the traditional view, partially unfolded proteins or unstructured regions of proteins are surprisingly common and are thought to be functionally important (*51, 53*). The highly flexible nature of the unstructured polypeptides enables them to bind to a variety of partners and facilitates participation in many biological processes. Many important proteins are either intrinsically unfolded or contain long segments of disordered regions such as the kinase-inducible activation (KIX) domain (*65-66*) of the transcriptional coactivator cyclic-AMP-response-element binding protein (CREB), cyclin-dependent kinase inhibitor p21 (*67*), the translation activation domain of protein p53 (*68*), and the ribosomal protein L11-C76 (*69*).

1.2.2 Characteristics of natively unfolded proteins

Many attempts have been made to predict protein disorder from the primary amino acid sequence. It has been established that many different disordered proteins share some common sequence features. For instance, the presence of low sequence complexity (*70-72*), defined by Wootton, is probably the signature of natively unfolded proteins. In addition, amino acid compositional bias, with low hydrophobicity and high polarity and net charge, correlates with intrinsic disorder (*52*). A number of algorithms are now available for prediction of disordered regions and the possibility of aggregation from amino acid sequences, including PONDR (*73*), TANGO (*74*), PASTA (*75*) and BETASCAN (*76*).

In general, unfolded polypeptide chains are relatively flexible and do not sufficiently bury hydrophobic residues to fold into well defined three dimensional structure. The unfolded state ensemble has a large distribution of structures and a wide range of different unfolded states can be populated under different conditions. Unfolded proteins can be expanded and behave as a random coils, or they can be relatively compact and contain significant secondary structure such as the molten globule state (*77*). Hydrophobic and aromatic clusters are also common in the unfolded states.

The formation of unfolded states can be driven by a variety of experimental conditions including high concentrations of denaturant, extremes of pH or high temperatures. It is worth noting that the unfolded state ensemble strongly depends on the

solvent conditions. The unfolded state under harsh denaturing conditions may be quite different from unfolded states under physiological conditions (*78-79*). Obviously the unfolded state that is formed under native conditions is the most physiologically relevant and also the most interesting for study. Unfortunately, the free energy barrier of protein folding and the high cooperativity of folding cause the unfolded state to be only sparsely populated at equilibrium under native conditions, preventing such states from being studied directly. Indirect approaches need to be applied in most cases.

1.2.3 Experimental methods to address unfolded proteins

Because unfolded proteins do not form crystals, X-ray crystallography is not applicable for the study of unfolded proteins. Circular dichroism and fluorescence spectroscopy, which are both sensitive to protein conformation, can provide valuable information about secondary structure and the local environment of fluorescent residues in unfolded proteins. As a result, they are both widely used to monitor unfolding. However they do not give residue specific information.

There is usually a significant change in the dimension of a protein as the folding/unfolding transitions occur, thus small angle x-ray scattering (SAXS) experiments, which measure the radius of gyration (*80*) (R_g) and pulsed field gradient (PFG) NMR diffusion experiments (*79, 81*), which measure the radius of hydration (R_h) can be used to characterize the overall average shape. NMR diffusion experiments are particularly useful because they do not require access to a synchrotron and can be easily applied to small proteins. In a NMR diffusion experiment, the intensity of a well-resolved resonance is monitored as a function of an applied magnetic field gradient and the data is fit to obtain the diffusion coefficient. The hydrodynamic radii are calculated based on the diffusion coefficient of the protein and an internal standard, typically dioxane.

Recent advances in NMR technology, including the emergence of ultra high-field instruments, the development of multidimensional heteronuclear experiments (*82-83*) as well as molecular biology methods for production of isotopically labeled proteins (*84-85*), have opened the door to characterize the unfolded state of proteins at atomic resolution. If the resonance assignments of the unfolded proteins can be obtained, NMR can provide much valuable information about the structural and dynamic properties of unfolded proteins using parameters such as chemical shifts (*86-87*), NOEs (*88*), relaxation rate constants (*89-91*), and residual dipolar couplings (*92-93*).

1.3 Protein cold denaturation

The phenomenon of protein cold denaturation has been observed for decades. For example, the rate of denaturation of the protein ovalbumin in concentrated urea was found to become higher at decreasing temperatures by Hopkins (*94*) and Clark (*95*). After that initial study, many enzymes were found to be inactivated at low temperatures in aqueous buffer solutions (*96-97*). Since decreasing the temperature should lead to a macroscopic system with increasing order and the native state of proteins are believed to be more ordered, protein cold denaturation can not be explained by classical thermodynamic principles. Motivated by the apparent paradox, protein cold denaturation has been extensively studied. Cold denaturation is now widely accepted to be a very general property of globular proteins. Similar to the denaturation driven by other

methods, the cold denaturatoin of small single-domain proteins is also usually highly cooperative (*98*).

Recent studies by Wand and coworkers, however, have led to the proposal that cold denaturation of proteins is highly non-cooperative and proceeds in a stepwise manner. Furthermore, these studies were suggested to provide a unique view of potentially important partially folded states. The work described in this thesis strongly argues that these conclusions are incorrect and result from problems in experimental design because the Wand experiments involved encapsulation in reverse micelles. Cold denaturation is largely due to the fact that the unfolded states of proteins are much more exposed to the solvent than the native state and the hydration of the non-polar hydrophobic groups is associated with a significance increase in the order of water and a decrease in the entropy of the system. Indeed, hydrophobic interactions, which are one of the major driving forces of protein folding, decrease at low temperatures (*98-99*). For example, the Gibbs free energy of transferring a non-polar substance from pure liquid to water, a measurement of hydrophobic interactions, has been shown to reach a maximum at a certain temperature for many non-polar organic compounds (*98, 100*).

According to the Gibbs-Helmholtz equation, the free energy of folding should be expressed by a parabolic function of temperature, provided that the heat capacity change upon the unfolding is not zero and is not strongly temperature dependent, which are reasonable approximations over the normal temperature range for proteins (*98*). As shown in Figure 1-3, there is a temperature at which the native state is the most thermodynamically stable. The native state becomes less stable at either higher or lower temperatures. As shown in Figure 1-3, there are two temperatures at which the free energy change ΔG^o (T) equals zero, the heat denaturation midpoint, T_m and a cold denaturation midpoint, T_C. Tc can be estimated using the following equation (*98*):

$$T_c = \frac{T_m^{\,2}}{T_m + 2(\Delta H^o(T_m)/\Delta C_p^{\,o})} \tag{8}$$

Where $\Delta H^o(T_m)$ is the enthalpy change at T_m, and $\Delta C_p^{\,o}$ is the heat capacity change between the native and denatured states. While the thermal unfolded state can be accessed at elevated temperature, the study of cold denatured state is often hampered because, for most proteins, T_C is well below the freezing point of water. This obstacle can be circumvented by using high pressure, adding a moderate concentration of denaturant (*55*), working with supercooled water (*101*), or encapsulation of proteins in reverse micelles (*36*). However, these artificial conditions may perturb the cold unfolded state of the proteins studied. An alternative approach is to find a protein or near native solvent conditions where cold denaturation can take place at temperatures above 0 ºC. In this thesis, the cold denaturation of a protein is achieved under native like conditions and is investigated in detail.

1.4 The ribosomal protein L9

1.4.1 The structure and function of the ribosomal protein L9

The unfolded state and transition state of the C-terminal domain of the ribosomal protein L9 are the focus of this dissertation. The ribosomal protein L9, a 149 residue protein, is a structural protein in the ribosome. The variant studied here is from *bacillus stearothermophilus (102)*. The crystal structure of protein L9 has been solved by Hoffman in 1996 (*103*). As shown in Figure 1-4, the protein has an unusual structural

feature: it is composed of two globular domains which are connected by a long central helix linker (residues 41-74). The helix is solvent exposed in the middle and both ends participate in the hydrophobic interactions in the two sub-domains. The protein does not contain any disulfide bonds and it does not bind to any metal ions or any cofactors. Both sub-domains of protein L9 have been found to fold independently in a two-state fashion (*33, 104*).

A 3.5 Å resolution crystal structure of the ribosome has been solved and the position of the protein L9 within the ribosome is found to be in the large subunit and close to protein L1, L2, L16, L19 and L28 (*105-106*). Each domain of the protein L9 binds to the domain V of the 23s rRNA and thus the entire protein acts as a molecular strut to stabilize the ribosome (*107*). L9 is highly soluble in water and expresses extremely well in E *coli*, with a typical yield of 70-80 mg per liter of LB media.

1.4.2 The C-terminal domain of the ribosomal protein L9

The C terminal domain contains 92 residues (residues 58-149 of protein L9) and consists of two loops, an α-helix and an unusual mixed parallel, anti-parallel three-stranded β-sheet packed against the central α-helix (*103*) (Ribbon diagram shown in Figure 1-5). The mixed parallel and anti-parallel β-sheet, which is an unusual structural feature, makes CTL9 an interesting target for folding studies. Previous studies in our group have shown that CTL9 folds in a two-state fashion (*32*). Both the folding rate and stability of CTL9 are strongly pH dependent because the acid unfolding process is coupled with the protonation of two buried histidine residues (His106 and His134). The protein is much more stable and folds more quickly at pH values above the native state pKa of the histidines. Acid-induced denaturation experiments conducted at room temperature revealed that CTL9 is well folded above pH 5.0 and fully unfolded below pH 2.5 in the absence of denaturant. Both the folded and unfolded states are populated significantly at the midpoint of acid-induced unfolding, which is around pH 3.8 and the acid induced unfolding is two state. Thus it is possible to significantly populate both the folded and unfolded state at pH 3.8 in the absence of denaturant and study their structural and dynamic properties under identical solvent conditions. Another interesting feature is that a single point mutation in the hydrophobic core of CTL9, the isoleucine 98 to alanine mutation, dramatically destabilizes the protein by 4 $kcal.mol^{-1}$ and leads to cold denaturation at temperatures above 0 ºC (*108*).

Measurement of the hydrodynamic properties of various unfolded states of CTL9 using PFG-diffusion experiments revealed that the acid induced unfolded states of CTL9 experience a significant expansion from pD 4.2 to pD 2.1, and the acid unfolded state of CTL9 at pD 2.1 has a very similar radius of hydration to the urea unfolded state of CTL9 at pD 3.8 (*79*). Here pD is defined as the apparent pH of a D_2O solution measured using a standard glass electrode, in other words, the uncorrected pH meter reading. The expansion of the acid unfolded state at low pH is consistent with the thermodynamic studies which suggest that both the m-values and the change of heat capacity change as the pH decreases, and the expansion is expected to be accompanied with some structural changes (*109*). The cold unfolded state was found to be relatively compact, however, the cold unfolded state displays a 20% increase in R_h as the temperature decreases from 25 ºC to 2 ºC (*108*).

1.5. The aims of this dissertation

The goals of the research described in this thesis are to characterize the structural and dynamic properties of the unfolded states of CTL9 at high resolution using NMR spectroscopy under various conditions, and to study the cooperativity of the cold denaturation, using the I98A mutant of CTL9 as a model protein.

NMR methods were used to study the pH 2.0 unfolded state and the urea unfolded state. Direct comparison of the pH 2.0 unfolded state with the urea unfolded state was performed to verify whether global parameters such as radius of hydration or radius of gyration are good indictors of protein random coil behavior. The unfolded state under more native like conditions, the pH 3.8 unfolded state, was studied under the same experimental conditions as the native state and was found to be significantly more structured than the pH 2.0 or urea unfolded state. The cold denatured state of I98A CTL9 was characterized in detail using NMR, at pH 5.7 and 12 ºC, in the absence of chemical denaturant. The cooperativity of the cold denaturation was tested.

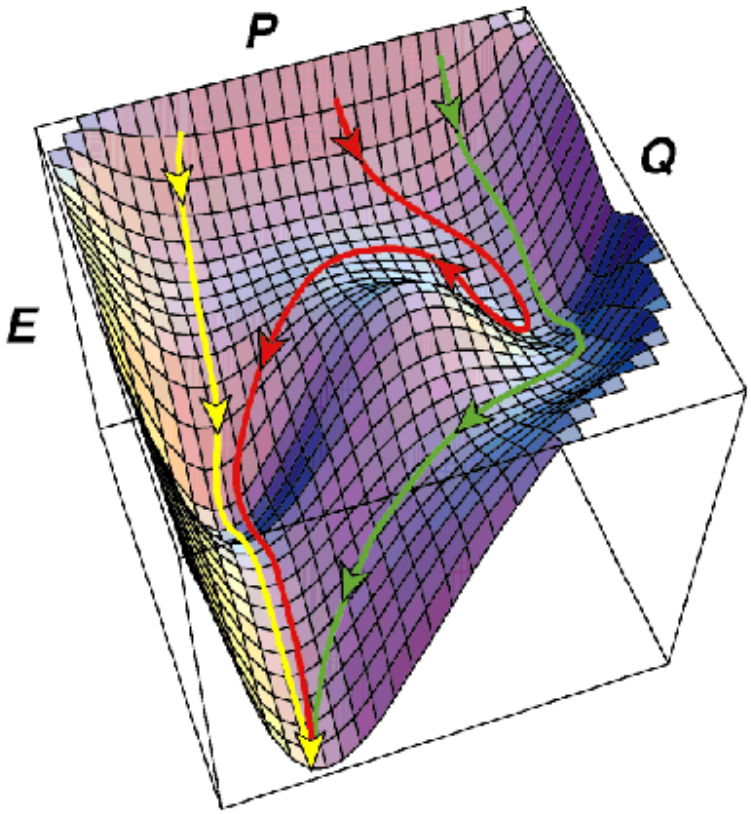


Figure 1-1. Three dimensional representation of a folding funnel. E represents the energy of the system, P is a measure of available conformational states and Q is defined as the proportion of native contacts formed. Adapted from ref (*26*).

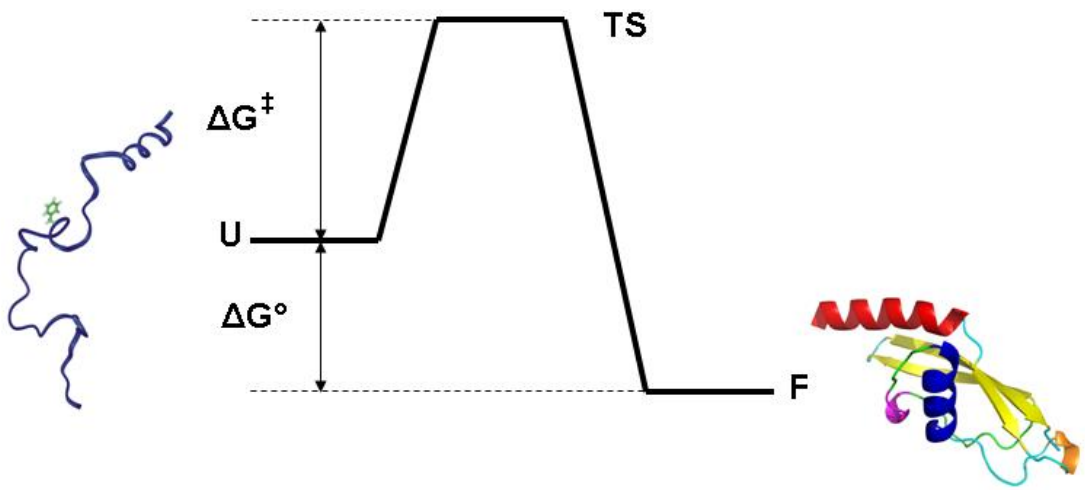


Figure 1-2. An energy diagram representing a two-state protein folding reaction. U, TS, F stand for the unfolded state, the transition state and the folded state respectively. ΔG^{o} represents the equilibrium free energy of folding and $\Delta G^{\neq}$ represents the activation free energy for folding.

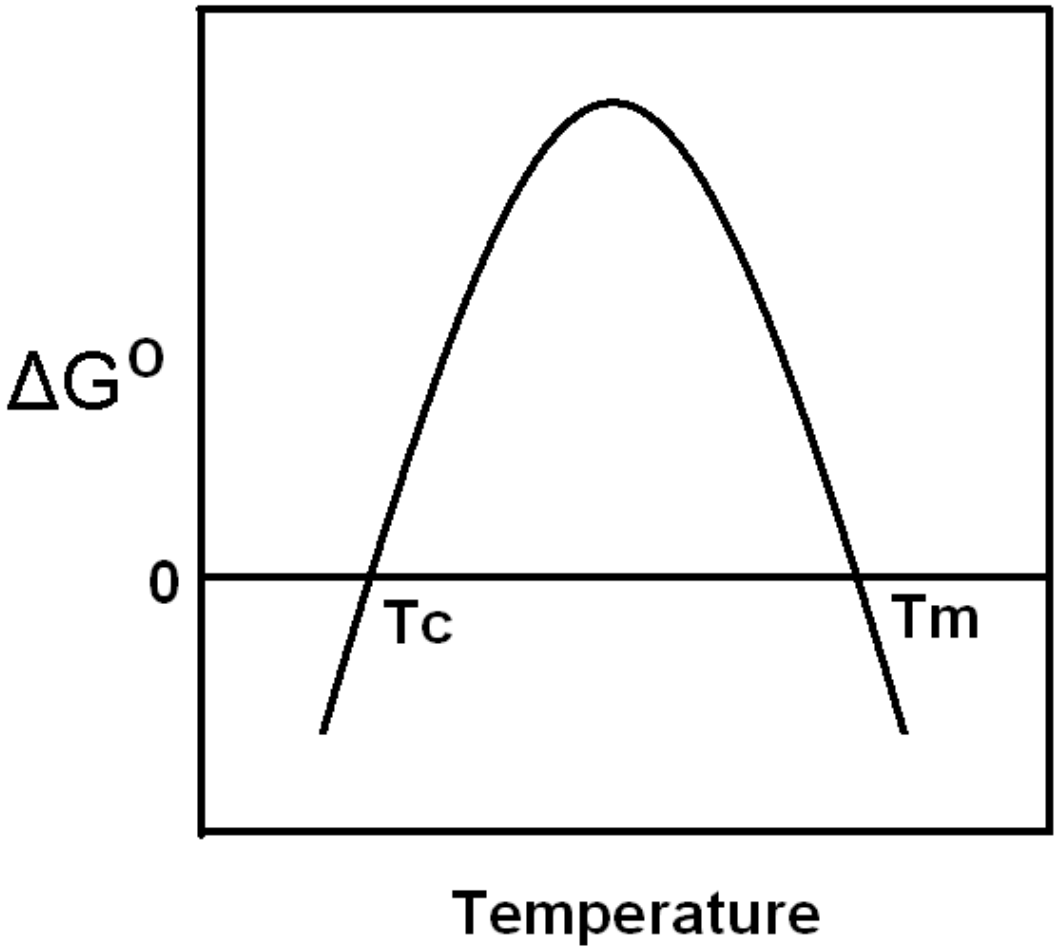


Figure 1-3. Plot of the free energy change of the protein unfolding reaction versus temperature.

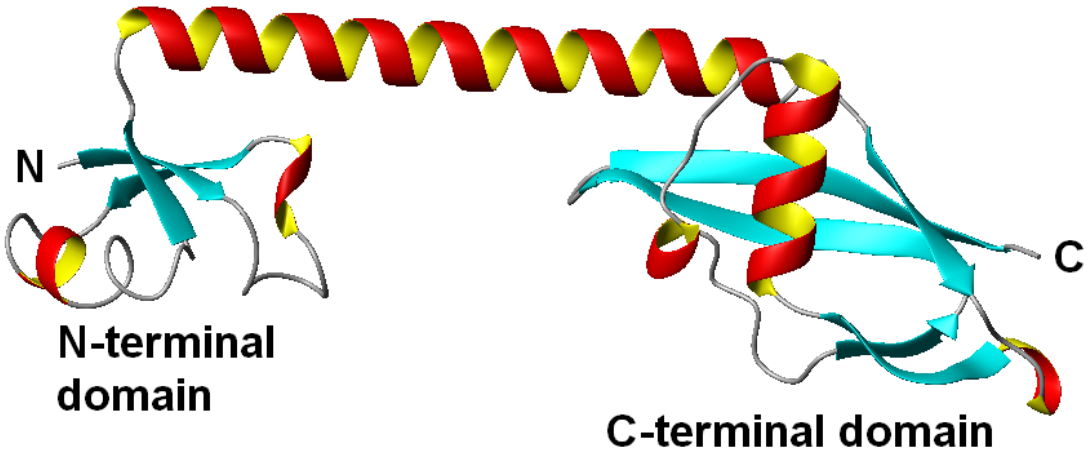


Figure 1-4. A ribbon diagram of the ribosomal protein L9. The PDB file 1DIV was used. The N and C termini are labeled.

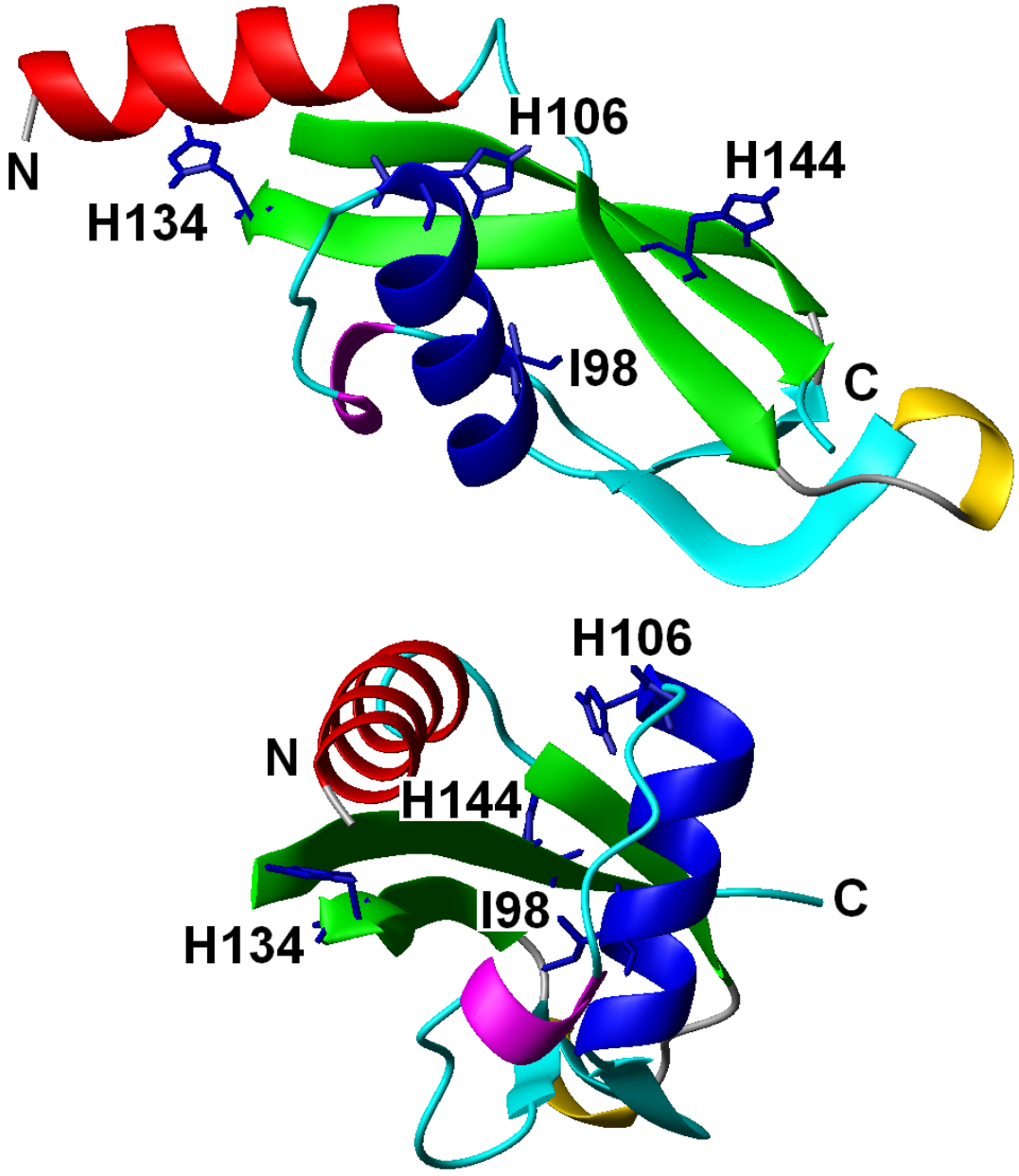


Figure 1-5. Two ribbon diagrams of CTL9 displayed in different orientation after rotation by 90^{o} about the z-axis. Side-chains of the three histidines, the hydrophobic core residue I98 and the two termini are labeled. The PDB file 1DIV was used.

2. The low pH unfolded state of the C-terminal domain of the ribosomal protein L9 contains significant secondary structure in the absence of denaturant but is no more compact than the low pH urea unfolded state

Abstract

There is considerable interest in the properties of the unfolded states of proteins, particularly unfolded states which can be populated in the absence of high concentrations of denaturants. Interest in the unfolded state ensemble reflects the fact that it is the starting point for protein folding as well as the reference state for protein stability studies, and can be the starting state for pathological aggregation. The unfolded state of the C-terminal domain (residues 58 to 149) of the ribosomal protein L9 (CTL9) can be populated in the absence of denaturant at low pH. CTL9 is a 92 residue globular α, β protein. The low pH unfolded state contains more secondary structure than low pH urea unfolded state but it is not a molten globule. Backbone (^{1}H, ^{13}C and ^{15}N) NMR assignments as well as side chain $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments and ^{15}N R_2 values were obtained for the pH 2.0 unfolded form of CTL9 and for the urea unfolded state at pH 2.5. Analysis of the deviations of the chemical shifts from random coil values indicates that residues that comprise the two helices in the native state show a clear preference to adopt helical φ, ψ angles in the pH 2.0 unfolded state. There is a less pronounced but nevertheless clear tendency for residues 107 to 124 to preferentially populate helical φ, ψ angles in the unfolded state. The urea unfolded state has no detectable tendency to populate any type of secondary structure even though it is as compact as the pH 2.0 unfolded state. Paramagnetic relaxation enhancement experiments indicate that the C-terminus of α-helix 1 forms long range interactions with the loop connecting α-helix 2 and β-strand 2 in the pH 2.0 unfolded state while there is no long range contact formed in the urea unfolded state. Comparison of the two unfolded forms of CTL9 provides direct experimental evidence that states which differ significantly in their residual structure can have identical hydrodynamic properties. This in turn demonstrates that global parameters such as R_h or R_g are very poor indicators of "random coil" behavior.

Note: The data presented here has been published (Shan, B., Bhattacharya, S., Eliezer, D., and Raleigh, D.P. *Biochemistry*, 2008, **47**, 9565-9573). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Prof. Raleigh and Prof. Eliezer. The PFG-NMR diffusion experiments were performed by Dr. Ying Li.

2.1 Introduction

Characterization of the unfolded state ensemble of proteins is an essential step in the development of a complete description of the protein folding process (*51, 110*). A detailed understanding of the unfolded state, particularly the nature and energetics of any residual interactions, is important because the unfolded state is the thermodynamic reference state for protein engineering studies, as well as the starting state for protein folding, and can be the starting state for pathological protein aggregation. Along these lines, recent work has highlighted the fact that mutations can exert significant effects upon the energetics of the unfolded state ensemble and can also alter a protein's propensity to aggregate (*54, 91, 111-113*). Interest in the unfolded state has increased dramatically with the realization that actual unfolded state ensembles can, and often do, deviate significantly from the classic random coil model (*54, 91, 111, 113-119*). It is now recognized that the properties and structural propensities of unfolded state can vary widely depending upon experimental conditions, ranging from relatively expanded ensembles with little propensity to form persistent structure to more compact ensembles containing elements of native and non-native structure (*120-123*). The latter are often found under more native conditions while the former are often populated under denaturing conditions. The unfolded ensemble populated under native conditions is clearly the most physiologically relevant state but it is difficult to characterize since it is normally sparsely populated owing to the cooperativity of folding and to the fact that the free energy balance favors the native state. These considerations have lead to increased interest in proteins whose unfolded states can be accessed in the absence of denaturant.

There is also considerable interest in characterizing the conformational propensities of the unfolded state ensemble at the level of individual residues. One question of particular importance is whether or not unfolded state structures are limited to native interactions or whether they include contributions from non-native interactions. A second important issue is whether or not global parameters which report on compactness such as the radius of hydration (R_h) or the radius of gyration (R_g) can be used to test if the unfolded state conforms to the random coil model. This is important because it is well known that R_h and R_g should obey a characteristic power law dependence on the number of residues in a random coil chain. However, computational studies have suggested that proteins can obey the "random coil" power law yet contain significant secondary structure (*114, 120, 122, 124-125*). Thus R_h and R_g may not be good probes of deviations from random coil behavior.

NMR is ideally suited for investigations of the unfolded state since it is currently the only method which can provide residue specific information for every residue in the unfolded protein (*83, 121, 123, 126-127*) The assignment of the NMR spectra of unfolded states can still be challenging because of limited spectral resolution. Fortunately, modern triple resonance methods can often overcome these problems and have opened the door to a detailed characterization of the unfolded state (*128*). Here we report the use of NMR to characterize the unfolded state of CTL9 in the absence of denaturant, and in the presence of 7.6 M urea. CTL9 is interesting in this regard since the low pH unfolded state in the absence of urea is no more compact than the urea unfolded state as judged by their respective radii of hydration.

CTL9 is a 92 residue globular protein derived from the ribosomal protein L9 (Figure 2-1). The protein adopts an interesting fold made up of two α-helices and an unusual three-stranded mixed parallel, anti-parallel β-sheet. The folding kinetics and equilibrium unfolding thermodynamics have been analyzed and equilibrium unfolding is well described as two-state (*32, 129*). The domain can be unfolded by heat, by denaturant or by low pH. The acid induced unfolded state contains residual secondary structure as judged by CD but it is not a molten globule (*79*). The goal of the present work is to compare the secondary structure propensities of an unfolded state of CTL9 that is populated in the absence of denaturant and the one which is populated under strongly denaturing conditions. The terms unfolded state, denatured state ensemble and denatured state are all found in the literature and there is no convention as to when a particular terminology should be used. We use the term pH 2.0 unfolded state here to refer to the unfolded state populated in the absence of denaturant at pH 2.0.

2.2. Materials and Methods

2.2.1. Mutagenesis, protein expression and purification

Primers were ordered from Operon and the plasmids for generation of the mutants were made and amplified using polymerase chain reaction (PCR). Mutations were validated by DNA sequence. Non-labeled proteins were expressed in *Luria-Bertani* media. Uniformly ^{15}N-labeled and ^{13}C, ^{15}N-labeled CTL9 and its mutants were expressed in *Escherichia coli* BL21 cells in M9 minimal media. The media for expressing ^{15}N-labeled protein contains 0.8 g/L $^{15}NH_4Cl$ as the sole nitrogen source and 10 g/L glucose as the sole carbon source, while 0.8 g/L $^{15}NH_4Cl$ and 10 g/L ^{13}C-glucose were used in the expression of ^{13}C, ^{15}N-labeled CTL9. Ampicillin was added to the main media at a concentration of 100 mg/L. The cells were grown at 37 ^{o}C until the optical density (O.D.) at 600 nm reached 0.7, at which point 100 mg/L IPTG was added to induce expression. After 4 hours, cells were harvested and then lysed by sonication. The protein was purified as described previously (*32*). The molecular mass of the protein was confirmed by mass spectrometry.

2.2.2. NMR sample preparation

Protein samples for NMR experiments were prepared in 10% D_2O at pH 2.0, no urea or at pH 2.5, in 7.6 M urea, both at a concentration of about 1 mM. The urea concentration was determined by refractometry. The sample used for the native state assignments was prepared in 20 mM sodium acetate and 100 mM NaCl at pH 3.8.

2.2.3. NMR spectroscopy

PFG-NMR diffusion experiments were performed on deuterated protein sample as previously described (*79*). All heteronuclear NMR experiments were performed on uniformly ^{15}N-labeled or ^{13}C, ^{15}N-labeled protein samples. All NMR spectra were recorded on a 700 or an 800 MHz Bruker spectrometer with a cryoprobe at the New York Structural Biology Center at 25 ^{o}C. In all NMR experiments, the ^{1}H dimension was centered at the water resonance and the ^{15}N offset frequency was set to 118.0 ppm.

^{15}N-^{1}H correlated heteronuclear single coherence (HSQC) spectra were collected using 1024 x 256 complex points and with 8 scans per increment. The spectra widths of the pH 2.0 sample were 7183.9 Hz and 2027.2 Hz for the ^{1}H and ^{15}N dimensions,

respectively. The pH 2.5, in 7.6 M urea sample had spectra widths of 6009.6 Hz (^{1}H) and 1561.0 Hz (^{15}N).

The following set of triple resonance experiments performed on a ^{13}C, ^{15}N-labeled sample of CTL9 were used to generate backbone assignments: HNCO, HN(CA)CO, HNCACB and CBCACONH. In the HNCACB and CBCACONH experiments, the ^{13}C offsets were 39.0 ppm, whereas the ^{13}C offsets were set at 176.0 ppm in the HNCO and HN(CA)CO experiments. For the pH 2.0 unfolded CTL9, the HNCACB and CBCACONH spectra were recorded with 1024 x 80 x 160 complex points, with spectra widths of 7183.9 Hz, 2027.2 Hz and 13076.2 Hz, in the ^{1}H, ^{15}N, and ^{13}C dimensions, respectively. The HNCO and HN(CA)CO spectra were acquired with 1024(^{1}H) x 72(^{15}N) x 144(^{13}C) complex points, with spectra widths of 7183.9 Hz (^{1}H), 2027.2 Hz (^{15}N) and 3018.4 Hz (^{13}C). For the pH 2.5, 7.6 M urea unfolded CTL9, the datasets for the HNCACB and CBCACONH spectra comprised 1024(^{1}H) x 72(^{15}N) x 128(^{13}C) complex points. The spectra widths were 6009.6 Hz, 1561.0 Hz and 13210.04 Hz in the ^{1}H, ^{15}N and ^{13}C dimensions, respectively. The HNCO and HN(CA)CO spectra were acquired with 1024(^{1}H) x 72(^{15}N) x 128(^{13}C) complex points, with spectra widths of 6009.6 Hz (^{1}H), 1561.0 Hz (^{15}N) and 2816.9 Hz (^{13}C).

3D HBHACONH experiments were carried out to determine the $^{13}C_{\alpha}$ ^{1}H and $^{13}C_{\beta}$ ^{1}H chemical shifts. For the pH 2.0 unfolded sample, the HBHACONH spectrum was collected with 1024 x 80 x 160 complex points, with spectra widths of 7183.9 Hz (direct ^{1}H dimension), 2027.2 Hz (^{15}N) and 7200.72 Hz (indirect ^{1}H dimension). For the pH 2.5, 7.6 M urea unfolded sample, the spectra widths were 6009.6 Hz (direct ^{1}H dimension), 1561.0 Hz (^{15}N) and 7001.1 Hz (indirect ^{1}H dimension) with 1024 (direct ^{1}H dimension) x 72 (^{15}N) x 128 (indirect ^{1}H dimension) complex points, respectively.

A 3D TOCSY-HSQC experiment was performed on a ^{15}N-labeled sample for the pH 2.0 unfolded state. A mixing time of 75 ms was used. The spectrum was recorded using a data matrix of 1024(direct ^{1}H dimension) x 128(^{15}N) x 256(indirect ^{1}H dimension). Spectra widths were 8012.8 Hz in the direct ^{1}H dimension, 2432.9 Hz in the ^{15}N dimension and 8001.6 Hz in the indirect ^{1}H dimension.

2.2.4. Measurement of transverse relaxation rates

The relaxation experiments were carried out as described (*130*). The spectra were collected at 10 delay times: 16.32 ms, 32.64 ms, 48.96 ms, 65.28 ms, 81.60 ms, 97.92 ms, 114.24 ms, 130.56 ms, 146.88 ms and 163.20 ms. A 900 μs delay was inserted between ^{15}N 180^{o} pulses in the Carr Purcell Meiboom Gill (CPMG) pulse train. Repeat measurements were made at 32.64 ms, 48.96 ms, 97.92 ms and 163.2 ms to allow estimation of uncertainty. Each T2 experiment was acquired with 4 scans using 1024 x 256 complex points. The spectra widths were 9615.4 Hz (^{1}H) and 2027.7 Hz (^{15}N) for the pH 2.0 unfolded state, and they were 7183.9 Hz (^{1}H) and 2027.6 Hz (^{15}N) for the pH 2.5, 7.6 M urea unfolded state. Recycle delay of 3 S was used.

2.2.5. Paramagnetic relaxation enhancement

3 mg of mutant of CTL9 containing a single cysteine was dissolved in 600 uL of NMR buffer. The pH was adjusted to 2, and 5 uL of a 300 mM MTSL stock solution was added. The reactions were performed overnight at room temperature (Figure 2-2, panel A), and the sample was loaded onto a Sephadex G25 column and centrifuged at 1500 rpm

to remove the excessive MTSL. Then the protein solution was split into two equal aliquots (300 uL each). 300uL of NMR buffer was added to one of the aliquots, 285 uL of NMR buffer together with 15 uL of 100 mM TCEP stock solution were added to the other to reduce the oxidized protein to the diamagnetic form (Figure 2-2, panel B). For the experiments in 8M urea, urea was added before splitting the samples and the concentration of urea was determined by measuring the refractive index. The completion of the reaction was monitored using analytical HPLC. A water/acetonitrile (ACN) gradient with 0.1% trifluoroacetic acid (TFA) as counterion was used with a 1.5 % gradient increments per minute. The protein eluted at about 37% acetonitrile was further confirmed using MALDI-TOF spectrometer. ^{15}N-HSQC experiments were collected for both the paramagnetic form and diamagnetic form. Assignments for the mutants were made based on the assignments of the wild-type CTL9 and verified using 3D TOCSY-HSQC experiments. The resonance intensities in the HSQC spectra of both the paramagnetic and diamagnetic states were measured and their ratios were calculated.

To calculate the theoretical intensity ratio for a completely non-structured polypeptide, the distance between each residue and the spin labels was calculated based on a random coil model assuming a Gaussian distribution of the room-mean-square end-to-end distance (*131-132*).

$$< r^2 >= nl^2\left(\frac{1+\alpha}{1-\alpha} - \frac{2\alpha(1-\alpha^n)}{n(1-\alpha)^2}\right) \quad (1)$$

where r is the end-to-end distance between a given residue and the spin label site, n is the number of residues between residue *i* and the spin label site, *l* is the link length of the chain, taken to a value of 3.8 Å (*125*), and α is the cosine of the bond-angle supplements for the freely rotating chain model, which was taken as 0.8 based on experimentally determined estimates of statistical segment lengths in poly-L-alanine (*132*). The contribution of paramagnetic enhancement to the transverse relaxation rate, R_{2p} was calculated from equation 2:

$$R_{2p} = \frac{K}{r^6}\left(4\tau_c + \frac{3\tau_c}{1+\varpi_H{}^2\tau_c{}^2}\right) \quad (2)$$

where r is the distance between each residue and the spin label site, K is 1.23 x 10^{-32} cm^2s^{-2}, ω_H is the Larmor frequency of a proton and τ_c is the effective correlation time for CTL9, which was estimated to be 4 ns based on NMR relaxation measurements (*133*). Then, the peak intensity ratios between the paramagnetic and diamagnetic forms were calculated using equation 3:

$$\frac{I_p}{I_D} = \frac{R_{2D}\exp(-R_{2p}t)}{R_{2D}+R_{2p}} \quad (3)$$

where R_{2D} is the transverse relaxation rate in the diamagnetic form, which was set to 3 s^{-1}, the approximate average R_2 obtained using NMR for CTL9 and t is the duration of the INEPT delays, which is 4 ms in HSQC pulse sequence (*134*).

2.2.6. Date processing and analysis

All spectra were processed using NMRPipe (*135*) and chemical shifts assignments were made using NMRView (*136*). All chemical shifts were referenced to the absolute frequency of DSS at zero ppm. Sequence dependent corrections of the chemical shifts

were made using methods developed by Schwarzinger, Wright, Dyson and coworkers (*137*). Random coil values in acidic (pH 2.3) 8M urea (*138*) were used to calculate the secondary chemical shifts, except that the values used for Glu and Val $^{13}C_{\beta}$ random coil shifts are those reported in NMRView algorithm. SSP analysis was performed using the method of Forman-Kay as described (*139*). The software provided at their website at http://pound.med.utoronto.ca/software.html was used. R_2 relaxation rates were determined using the automated program in NMRView by fitting the peak intensities to equation 4 which describes a two-parameter exponential decay:

$$I(t) = I_0 \exp(-t/T_2) \quad (4)$$

where $I(t)$ is the peak intensity after a delay of time t and I_0 is the intensity at the time point t=0. The R_2 rates were analyzed using a simple model by fitting the experimental R_2 rates to equation 5:

$$R_2(i) = R_2(\text{int}) \sum_{j=1}^{N} \exp\left(-\frac{|i-j|}{\lambda}\right) \quad (5)$$

where $R_2(i)$ is the experimental R_2 value for residue i, $R_2(\text{int})$ is the intrinsic relaxation rate which depends on the temperature and the viscosity of the solution, λ is the persistence length of the chain and N is the total number of residue in the protein (*91*).

2.3. Results and Discussion

2.3.1. Hydrodynamic properties of the unfolded states of CTL9

The hydrodynamic properties of the acid unfolded states and urea unfolded states populated under various pH values were measured using PFG-NMR diffusion experiments. The hydrodynamic radii are summarized in Table 2.1. The experimental measured R_h is generally consistent with predicted values based on empirical correlations (*81*), however, a large range of R_h was observed for the acid unfolded state. The R_h of the acid unfolded states populated in absence of urea increases from 25.1 Å to 33.5 Å as the pD decreases from 4.2 to 2.1. The apparent expansion of the acid unfolded state may be explained by the increasing net charge of the protein at lower pH. It is also likely associated with the decrease in the amount of residual secondary structural elements in the unfolded states. The unfolded states of CTL9 populated in 8M urea are generally more expanded than the acid unfolded states. Nevertheless, the pD 2.1 unfolded state without urea has a very similar hydrodynamic radii as the pD 3.8, 8M urea unfolded state. This observation seems to be contradictory with the CD experiments, which indicate that the acid unfolded state of CTL9 at pH 2.5 contains more secondary structure than the urea unfolded state at both pH 2.5 and pH 3.8. However, this also provides an excellent experimental example to test whether or not hydrodynamic parameters such as R_g or R_h may be accurate predictors of random coil behavior for unfolded proteins. As a result, the structural properties of the acid unfolded states and urea unfolded states under various pH values were characterized and compared in detail.

2.3.2. Sequence specific assignments of CTL9 in the pH 2.0 acid unfolded state and in the pH 2.5, 7.6 M urea unfolded state

The unfolded state of CTL9 populated at pH 2.0 in the absence of urea and the pH 2.5, 7.6 M urea unfolded state were characterized using ^{15}N, ^{1}H, ^{13}C triple resonance methods. We choose pH 2.5 for the urea unfolded state studies since low pH solutions of

high urea concentration require the addition of significant amounts of acid owning to the buffering capability of urea near pH 2.0. Thus adjusting the pH to 2.0 would lead to a noticeable increase in the ionic strength. The pKa's of all of the acidic residues in the urea unfolded state are expected to be very close to model compound values thus the choice of pH 2.0 or pH 2.5 should not affect the ionization state of any titratable side chains for the high urea studies. The HSQC spectrum of the pH 2.0 unfolded state of CTL9 is typical of that expected for an unfolded protein (Figure 2-3, panel A). A number of clearly resolved resonances are observed but a large number of peaks are clustered together in a limited frequency range, spanning less than 1 ppm in the ^{1}H domain but covering about 20 ppm in the ^{15}N dimension. The limited dispersion is typical of unfolded proteins but is worse in the case of CTL9 because of the small number of aromatic residues in the sequence. The peaks are sharp as expected for a monomeric protein. Previously reported PFG-NMR diffusion experiments have shown that CTL9 is monomeric under the conditions of these studies (*79*). The HSQC spectrum of the urea unfolded state is also poorly resolved (Figure 2-3, panel B), nevertheless, nearly complete ^{13}C, ^{1}H, ^{15}N backbone, $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments could be obtained. We also assigned the native state spectrum at pH 3.8 in order to provide a basis for comparison with the unfolded state data. The folded state assignments have been reported for CTL9 at pH 5.5. However, complete $^{13}C_{\beta}$ assignments are not available in the literature (*103*). In addition, pH 3.8 is closer to the pH used for the unfolded state studies.

HNCACB and CBCACONH triple resonance experiments were used to establish backbone connectivities for both states. The HNCO and HNCACO experiments were used to confirm these assignments and obtain ^{13}C carbonyl assignments. $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ chemical shifts were obtained from analysis of the HBHACONH spectrum while $^{13}C_{\beta}$ assignments were determined using the HNCACB experiment. HSQC-TOCSY experiments were used to confirm the assignments and aided in the assignments of resonances in residues immediately proceeding prolines. Complete backbone assignments were determined for 89 out of 92 residues, in the pH 2.0 unfolded state. The sole exceptions are the three prolines (P80, P130 and P135), for which we obtained ^{13}C carbonyl, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ assignments but not ^{15}N assignments. ^{15}N and amide ^{1}H assignments were obtained for more than 85 % of the non-proline residues in the urea unfolded state and $^{13}C_{\alpha}$, $^{1}H_{\alpha}$, $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments were obtained for all but 3 of the non-glycine residues, while ^{13}C carbonyl assignments were made for 89 out of the 92 residues. No peaks could be detected for L62, L102 and L117. Overall at least partial assignments were obtained for 89 out of 92 residues. Native state assignments at pH 3.8 were obtained using the same strategy and are used here to compare with our unfolded state assignments. All of the assignments have been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu/).

2.3.3. Analysis of ^{13}C, ^{1}H and ^{15}N chemical shifts provides evidence for native secondary structure in the unfolded state in the absence of denaturant

It is well known that secondary chemical shifts, i.e., the deviations between observed and random coil chemical shifts, are very sensitive to secondary structure. We analyzed the deviations between our experimentally determined chemical shifts and the random coil values reported for pH 2.3 in 8M urea (*138*), as reported in NMRView. Since these chemical shifts were measured at pH 2.3, no pH correction is needed. Sequence

dependent corrections of ^{13}C carbonyl chemical shifts were made using reported protocols (*137*).

Figure 2-4 shows plots of the deviations of the C_α ^1H chemical shifts. C_α ^1H shifts are sensitive to secondary structure and are shifted upfield in α-helices and downfield in β-sheets. The panel A of Figure 2-4 displays the deviations for the native state at pH 3.8, the second shows the plot for the pH 2.0 unfolded state, and the third panel displays the deviations for the urea unfolded state at pH 2.5. The deviations are large for the folded state, spanning the range of -0.73 ppm to 1.05 ppm. The secondary shifts of the folded state are consistent with the known secondary structure of the protein. As expected, the deviations are much smaller for the pH 2.0 unfolded state but are clearly different from zero. The sequences corresponding to the first and second α-helices display contiguous negative deviations. For α-helix-1, the average deviation is -0.05 ppm while it is -0.09 ppm for α-helix-2. For comparison, both of these regions in the native state show average deviations of -0.23 ppm. The deviation from random coil values obtained for the sequence corresponding to the β-sheet exhibit no particular trend. A consecutive set of negative deviations (-0.07 ppm in average) is also observed from residue 112 to 114, a segment forming a 3_{10} helix in the native state. The plot for the urea unfolded state (Figure 2-4C) displays even smaller negative deviations of Cα ^1H chemical shifts, as compared with the pH 2.0 unfolded state. The averages are -0.04 and -0.06 ppm for α-helix-1 and α-helix-2 regions, respectively, while the segment corresponding to the 3_{10} helix does not display any obvious trend. Taken in isolation, the Cα ^1H deviations suggest a tendency to populate helical regions of the φ, ψ plot in the pH 2.0 unfolded state for the residues which are helical in the native state. In the urea unfolded state, the tendency to form α-helical structure is much weaker. The data are suggestive, however, C_α ^1H secondary shifts are usually small for unfolded proteins and therefore may not be sensitive to small populations of residual secondary structure in a disordered protein. A more accurate and complete picture can be obtained by analyzing the $^{13}C_\alpha$, $^{13}C_\beta$, and ^{13}C carbonyl shifts.

Figure 2-5 displays plots of the deviations of the $^{13}C_\alpha$ (Figure 2-5A), carbonyl ^{13}C (Figure 2-5B), $^{13}C_\beta$ chemical shifts (Figure 2-5C) from random coil for the pH 2.0 unfolded state. $^{13}C_\alpha$ and ^{13}CO shifts are expected to exhibit positive secondary shifts in α-helices and negative secondary shifts in β-strands. The pattern of the $^{13}C_\alpha$ secondary shifts is consistent with the C_α ^1H secondary shifts. The largest deviations are found in the regions which correspond to the two native state α-helices. The average deviation for the segment corresponding to α-helix-1 is 0.32 ppm while it is 0.62 ppm for α-helix-2. Positive deviations are also observed for the loop region which connects α-helix-2 to β-strand-2 (residues 107 to 124), however they are somewhat smaller (0.19 ppm). No strong trend is observed for the β-sheet regions. A similar pattern was observed for the ^{13}CO secondary shifts. The average is 0.41 ppm for α-helix-1 and it is 0.70 ppm in α-helix-2. Smaller positive deviations (0.28 ppm) are observed in the loop region between α-helix-2 and β-strand-2, in agreement with the smaller $^{13}C_\alpha$ and $^1H_\alpha$ secondary shifts for the region.

$^{13}C_\beta$ secondary shifts are also sensitive to secondary structure but the sign of the deviation is reversed relative to that observed for $^{13}C_\alpha$ shifts (*87, 138*). The $^{13}C_\beta$ secondary shifts are consistent with the C_α ^1H, $^{13}C_\alpha$ and ^{13}CO secondary shifts. Negative secondary shifts are detected for the segments corresponding to two α-helices in the

native state, and for the connection between α-helix-2 and β-strand-2. The average values for α-helical regions are -0.26 ppm and -0.33 ppm, respectively, while the average for the connection region is -0.23 ppm. No consecutive regions of positive deviations are observed for segments corresponding to the three β-strands.

Figure 2-6 displays plots of the secondary shifts for the $^{13}C_{\alpha}$ (Figure 2-6A), carbonyl ^{13}C (Figure 2-6B), $^{13}C_{\beta}$ (Figure 2-6C) chemical shifts for the pH 2.5, 7.6 M urea unfolded state. No consecutive set of residues can be found to display large secondary shifts. This indicates that the presence of a high concentration of denaturant abolishes most of the residual structure in the pH 2.0 unfolded state, making the urea unfolded state very different from the pH 2.0 unfolded state in terms of the amount of residual secondary structure elements. Consideration of the individual plots of secondary shifts gives rise to a self-consistent picture. Residues corresponding to α-helix-1 and α-helix-2 have a tendency to preferentially populate the helical region of the Ramachandran plot in the pH 2.0 unfolded state in the absence of urea. Also there are suggestive hints for α-helical propensity in the sequence between α-helix-2 and β-strand-2. The results are quite different for the pH 2.5, urea unfolded state. The deviations from random coil values are much smaller and no clear trends are detected.

Individual secondary shifts are useful but, recent work has shown that combining these data can provide more reliable information. A common approach is to calculate the difference between $^{13}C_{\alpha}$ secondary shifts and $^{13}C_{\beta}$ secondary shifts. This method has the advantage of canceling out any uncertainty in the chemical shift reference. Positive $\Delta\delta C_{\alpha}$-$\Delta\delta C_{\beta}$ values indicate α-helical structure and negative values indicate propensity to adopt β-strand structure (*139*). Figure 2-7 displays the plot of the difference in $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ secondary shifts. Several obvious trends are readily apparent for the pH 2.0 unfolded state. Positive values are observed for α-helix-1 and α-helix-2. The average for α-helix-1 is 0.58 ppm and for α-helix-2 is 0.95 ppm. A set of consecutive positive deviations are observed for the loop between α-helix-2 and β-strand-2 as well. The average for this region, residue 107 to residue 124, is 0.44 ppm. The difference is much smaller for the urea unfolded state and no clear pattern is detected. The average deviation for helix-1 region is -0.10 ppm, while it is -0.02 ppm for helix-2 and only -0.09 ppm for residues 107 to 124.

Recently, Forman-Kay and coworkers have used the weighted average of the secondary shifts to probe secondary structure where the weighting factors reflect the relative sensitivity of the secondary shifts to the type of structure. The method, known as SSP analysis, has recently been applied to the synucleins (*139*). Figure 2-8 shows the results of the SSP analysis applied to the pH 2.0 unfolded state and to the urea unfolded state. In this analysis, the higher the value of the SSP score, the greater the tendency to adopt helical φ, ψ angles. A value of 1.0 is expected for a residue which adopts helical φ, ψ angles 100% of the time and a value of -1.0 for a residue which exclusively populates the β-sheet region. The analysis is based upon the assumption that an observed secondary shift, relative to the average secondary shift for fully formed secondary structure represents the fractional population of the structure at a particular site. As such, SSP scores should exhibit a correspondence with the propensity to adopt secondary structure. The regions corresponding to α-helix-1 and α-helix-2 clearly have the highest SSP scores for the pH 2.0 unfolded state, with average scores of 0.28±0.07 and 0.32±0.07, respectively. The value for the connection linking α-helix-2 and β-strand-2 is 0.19. On

the surface, the positive SSP score for this segment might imply that non-native structure is populated in this region. However, it is also possible that the native fold could generate positive SSP scores. Calculation of the SSP score for the pH 3.8 native state gives an average of 0.25 for residues 107 to 124, indicating that positive SSP values in this segment can result from native state structure, thus it is formally not clear, if the chemical shifts deviations between residues 107 to 124 represent a tendency to populate native or non-native structure (Figure 2-8). However, the key result is that the SSP analysis is completely consistent with all other measured parameters and indicates a tendency to form structure. The plot clearly reveals that helix propensity is significantly higher in the pH 2.0 unfolded state for regions of the protein that are helical in the native state. In contrast, the urea unfolded state shows no statistically significant tendency to populate non-random structure. The SSP values are -0.01±0.06 for α-helix-1, -0.05± 0.09 for α-helix-2 and -0.04± 0.07 for the connection between α-helix-2 and β-strand-2. Forman-Kay and colleagues have pointed out that ^{1}H C_{α} secondary shifts can sometimes be unreliable and discussed how their inclusion or exclusion can alter the SSP analysis (*139*). Consequently, we repeated the analysis without including the ^{1}H C_{α} secondary shifts. The overall shape of plot is essentially unaltered and the conclusions are not affected (Figure 2-9).

2.3.4. ^{15}N R_2 relaxation analysis of the low pH unfolded state and the low pH urea unfolded state of CTL9

^{15}N R_2 relaxation rates have been used to detect apparent deviations from random coil behavior in unfolded proteins. Schwalbe and coworkers have pioneered the approach (*90-91*). In their analysis, relaxation data is fit to a purely phenomenological model which assumes that the R_2 value at a particular residue is correlated with the values of R_2 at residues adjacent in primary sequence. The length of the correlated segment can, at some level, be viewed as being related to the persistence length. For a chain with only local interactions, a characteristic inverted U shape is predicted. Deviations from the predicted curve are typically taken to represent persistent clustering of hydrophobic side chains, although an atomic level description of the origin of the effect is not available. Figure 2-10 compares plots of the measured R_2 values and the R_2 values calculated using the random coil model. The calculated R_2 values follow an inverted U-shaped curve, showing uniform relaxation rates in the middle of the polypeptide chain, with an average of 2.74 s^{-1}. Smaller rates are observed at both terminal of the chain. There are no significant differences between the experimental data for the pH 2.0 unfolded state and the urea unfolded state. There are deviations from the phenomenological "random coil" model for both data sets, but these are smaller than what have been observed for other proteins (*63, 90*). The simplest interpretation of the relaxation data is that there are no long lived clusters in the unfolded state(s), of course one should bear in mind that the analysis involves comparison of experimental data to a phenomenological model.

2.3.5. Mapping long range interactions in the unfolded states of CTL9

Because of the highly flexible nature of the unfolded polypeptide chain and the severe resonance overlapping, medium range or long range NOEs were not observed for either the acid unfolded state or the urea unfolded state. Paramagnetic nitroxide spin labels lead to substantial resonance line broadening of the spins which are located with 15

Å of the site of the labeling, and thus are very useful probes of long range interactions in unfolded proteins (*131, 140-142*).

Wild-type CTL9 does not contain any cysteine residues. Thus site-directed mutagenesis was used to introduce single-cysteine mutants which are then modified with *S*-(2, 2, 5, 5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) spin label reagent, as illustrated in Figure 2-2, panel A. To avoid possible perturbation of the structure due to the mutation, three solvent exposed residues were selected as the sites to introduce the spin labels: K74 at the C-terminus of the first α-helix, K109 and D119 in the loop region between the α-helix 2 and β-strand 2. The resulting mutants, denoted as K74C, K109C and D119C have identical CD spectra as the wild-type within experimental error, indicating equivalent amounts of secondary structure content. The ^{15}N-HSQC spectrum of the diamagnetic form of each mutant was compared to the spectrum of the wild-type (Figure 2-11). Most peaks except for the residues immediately adjacent to the site of mutation do not shift their positions significantly or show significant line broadening, suggesting that the mutations do not change the structure of the protein.

^{15}N-HSQC spectra were collected for both the paramagnetic form and diamagnetic form of each mutant. The paramagnetic enhancement of the nuclear relaxation was studied as a function of primary amino acid sequence by measuring the ratios of HSQC peak intensities of the paramagnetic form over those of the diamagnetic form. A ratio of 1 indicates no detectable broadening due to the paramagnetic enhancement and smaller numbers indicate increasing paramagnetic effects. Figure 2-12 shows the paramagnetic relaxation enhancement for the urea unfolded states of the three mutants. For all three mutants, most of the cross peaks are not affected by the spin labels. Substantial broadening are only observed for those residues within ±5 amino acids of the site of the spin labels, suggesting that no long range interactions are formed in the urea unfolded state. The paramagnetic enhancement of the acid unfolded state in the absence of urea is plotted in Figure 2-13. The spin labels clearly have different effects on the relaxation enhancement for the acid unfolded state compared to the urea unfolded state. For K74C CTL9, the most drastic enhancement occurs within residues 72 to 79, where the resonances diminish due to the line broadening. Significant broadening are also observed throughout residues 80 to 136, indicating that residue K74, which is located at the C-terminus of α-helix 1, forms transient long range contact with those residues. Resonance broadening caused by spin labels at positions 109 and 119 confirm the interaction between the C-terminus of α-helix 1 and loop connecting α-helix 2 and β-strand 2. In addition, spin labels at positions 109 and 119 cause resonance broadening in α-helix 2 as well as β-strand 2, suggesting there might be transient long range contact formed between these segments. Introducing spin labels at more sites will further define the long range structure in the acid unfolded state of CTL9.

2.4. Discussion and conclusions

We have obtained essentially complete ^{1}H, ^{13}C and ^{15}N backbone assignments, as well as $^{13}C_{\beta}$ and ^{1}H C_{β} assignments for the pH 2.0 unfolded state of CTL9 and virtually complete assignments for the pH 2.5 urea unfolded state. The secondary shifts, the difference between the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ secondary shifts and the SSP analysis all yield a self consistent picture; namely that residues which are helical in the native state have a

propensity to preferentially populate helical φ, ψ angles in the pH 2.0 unfolded state. Residues 107 to 124, which form the connection between the second α-helix and the second β-strand, also exhibit a clear, albeit reduced, tendency to preferentially sample the helical region of the Ramachandran plot. Thus the unfolded state of CTL9 contains significant secondary structure in the absence of denaturant. There is a very good correlation between the regions predicted by the chemical shift analysis to sample helical φ, ψ angles and the regions predicted to have the highest helical propensity by the AGADIR algorithm of Serrano and coworkers (*143*). AGADIR predicts that the first and second α-helices and the loop connecting the second α-helix to the second β-sheet have the highest helical propensity in CTL9 (Figure 2-14). The urea unfolded state is very different and displays, as judged by the NMR analysis, no tendency to preferentially populate secondary structure.

It is difficult to precisely deduce the fractional population of a particular type of secondary structure from analysis of chemical shift data alone but the SSP method appears to be the most promising approach. It is thus interesting to compare the SSP analysis to estimates of the helical content based upon CD measurements. The mean residual ellipticity of the pH 2.0 unfolded state is -4900 deg $cm^{-1}mol^{-1}$ at 222 nm. The apparent fractional population of helix can be estimated from CD spectra if values for a fully helical peptide and for an unstructured peptide are known. The largest ambiguity arises in the choice of the ellipticity value for the random coil state. If we use the actual ellipticity measured in 8M urea, we obtain an estimate of 8.2 % α-helix. If, instead, we use empirical relations derived by Baldwin and coworkers, we obtain a value of 12.6 % α-helix (*144*). The SSP analysis leads to a somewhat higher predicted fraction of α-helix. The average SSP score for the entire sequence is 0.21, which, in the simplest interpretation corresponds to 21 % α-helix (a SSP score of 1.0 is taken to represent a 100% tendency to populate helical φ, ψ angles). The discrepancy may reflect difficulties associated with using the SSP methodology to precisely quantify the relative population of secondary structure. Of course, difficulties in interpreting the CD spectra of partially structured polypeptides could play a factor too. However it is worth bearing in mind that CD and chemical shifts have a fundamentally different sensitivity to helical structure and we believe this is the most likely explanation. Secondary shifts are largely local in origin and a short segment of chain that exclusively populates the helical region of φ, ψ space would give rise to SSP scores near 1.0. In contrast, the CD signal of α-helices is well known to depend on the length of the helix and short helices can give rise to much weaker intensity at 222 nm than longer helices (*145*). Thus the difference between the NMR and CD analysis could reflect a tendency for individual residues or short sections of the chain to adopt helical φ, ψ angles even though there is a much more modest propensity for a consecutive set of segment of residues to cooperatively sample helical conformations. Details aside, the NMR studies and CD measurements both demonstrate that the low pH unfolded state of CTL9 clearly deviates from the classic random coil model.

In addition to the difference in the secondary structure content, the low pH unfolded state and the urea unfolded state also differ in their ability to form long range contacts, as shown by the paramagnetic relaxation enhancement experiments. The low pH, urea unfolded state does not have any tendency to form detectable long range interactions. In comparison, there are significant long range contacts found in the low pH unfolded state

in the absence of urea. The most obvious long range interactions are observed between the C-terminus of α-helix 1 and the connecting loop between α-helix 2 and β-strand 2. To obtain a better picture of the global structural propensities, spin labels at more sites will be helpful. Three additional mutants have been produced and modified with MTSL, they are residues 61, 96 and 149. NMR analysis of these mutants is being carried out by Mr. Bowu Luan.

At first glance, the observation of helical structure in the low pH unfolded state may seem to be at variance with hydrodynamic measurements. Those studies showed the pD 2.1 unfolded state has the same radius of hydration, R_h = 33.5 Å, as the urea unfolded state at pD 3.8, R_h = 33.6 Å, even though the urea unfolded state has much less, if any, residual helical structure (*79*). In addition, the R_2 data collected here for the pH 2.0 unfolded state and the pH 2.5, urea unfolded state are broadly similar. The apparent discrepancy simply reflects the fact that global parameters such as R_h or R_g can be insensitive to the presence of even significant amounts of secondary structure. The spin label data show that they are also insensitive to transient long range contacts detected by the paramagnetic relaxation enhancement methods. This effect has been predicted based upon computational studies by Pande and Rose (*122, 125*). Those workers compared the calculated scattering profiles of random-flight chains to identical sequences that contained segments of α-helical structure. They found that the calculated values of R_g were very similar even though the sequences have quite different amounts of secondary structure. CTL9 provides a striking experimental example of this principle.

It is interesting to consider the results in light of a recently published analysis of the transition state of the folding of CTL9 (*129*). That study made use of φ-value approaches and found that the largest φ-values were localized in the β-hairpin encompassing β-strands two and three and the connecting loop. Small φ-values were observed in mutations in both helix-1 and helix-2. Small φ-values, in the absence of unfolded state effects, indicate that the region being probed is no more structured in the transition state than in the unfolded state. However, small φ-values do not necessarily mean that the region of the protein being studied is devoid of structure in the transition state. Instead, small φ-values could arise because the region is similarly structured in both the unfolded state and the transition state (*146*), thus there is a structural ambiguity in the interpretation of small φ-values. The analysis presented here indicates that the two helices are partially formed in the unfolded state and thus also partially structured in the transition state for folding.

Acknowledgements

We thank Prof. David Hoffman for providing the pH 5.5 native state assignments of CTL9 and we thank Dr. Ying Li for her contributions for the measurement of the hydrodynamic properties of the unfolded states of CTL9.

Table 2-1. Hydrodynamic radii for the native, acid unfolded and urea unfolded states of CTL9

	R_h (Å)	
pD[a]	unfolded protein	denatured protein in 8 M urea
2.1	33.5	
3.4	27.2	
3.8	28.1	33.6
4.0	25.1	
4.2	25.1	
4.7	-[b]	32.6
5.5		30.7
7.0		30.3

The standard error to the fit for each curve is on the order of 0.5 Å. However this will underestimate the experimental uncertainty. The uncertainty was estimated to be ±5% based on repeated measurements.

[a.] Uncorrected pH meter reading in D_2O.

[b.] Resolved resonances for the unfolded state were not observed above pD 4.2.

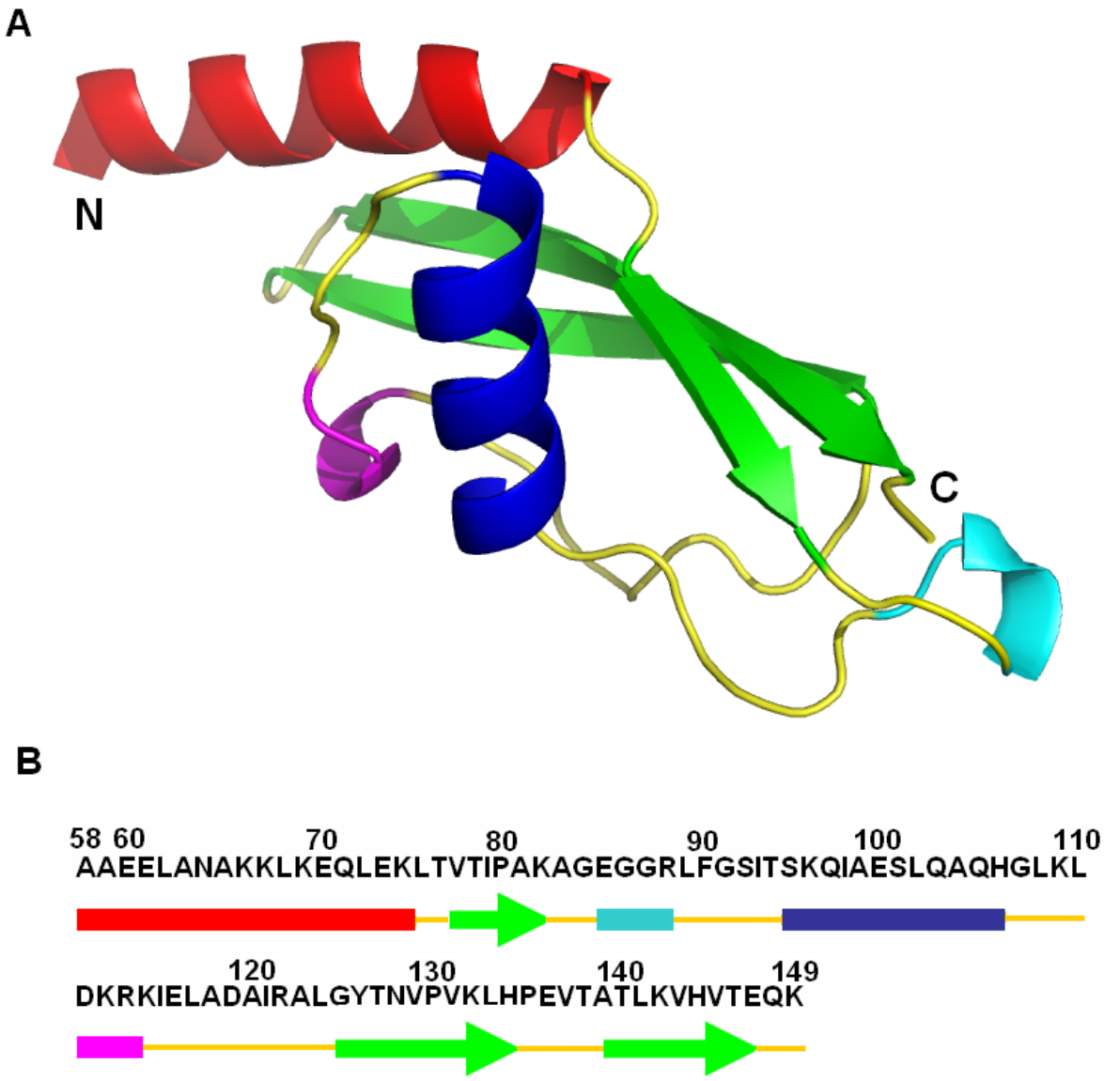


Figure 2-1. The structure of CTL9. (A) Ribbon diagram of residues 58 to 149 of L9 (pdb file code1DIV). The N and C termini are labeled. (B) The primary sequence of CTL9 is shown together with a schematic representation of the different elements of secondary structures (arrows represent β-strands, colored cylinders represent α-helices and 3_{10} helices). The ribbon diagram was constructed using PyMOL.

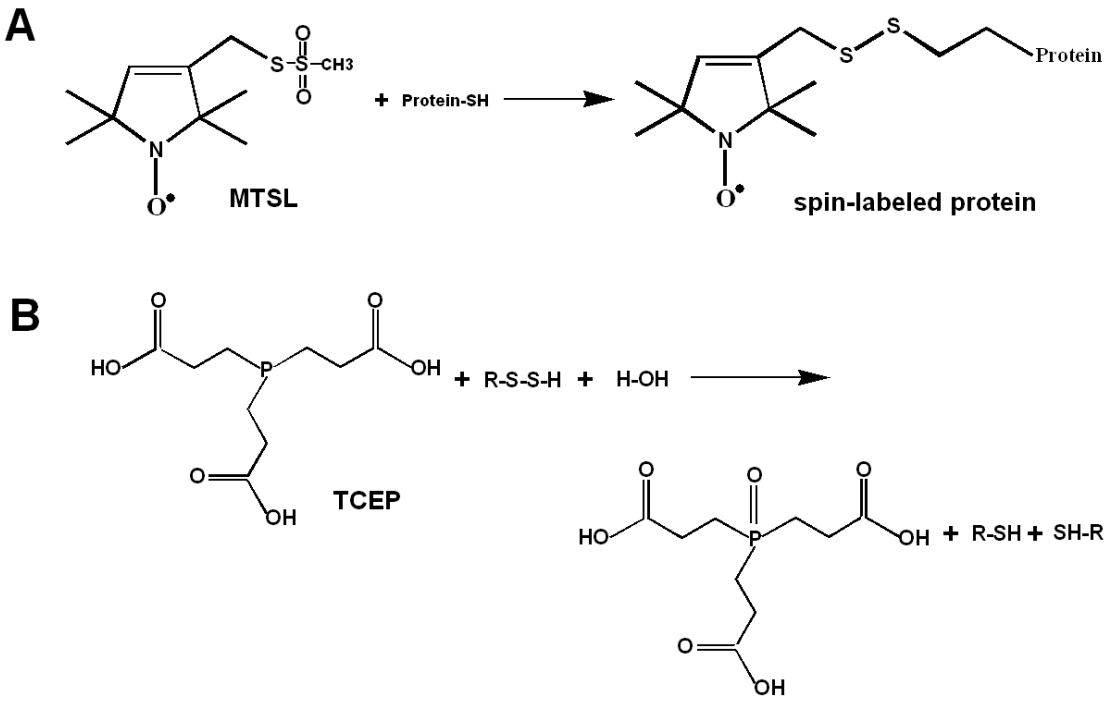


Figure 2-2. (A) Reaction scheme of MTSL with the cysteine sidechain of a protein. (B). Reaction scheme of TCEP with oxidized protein.

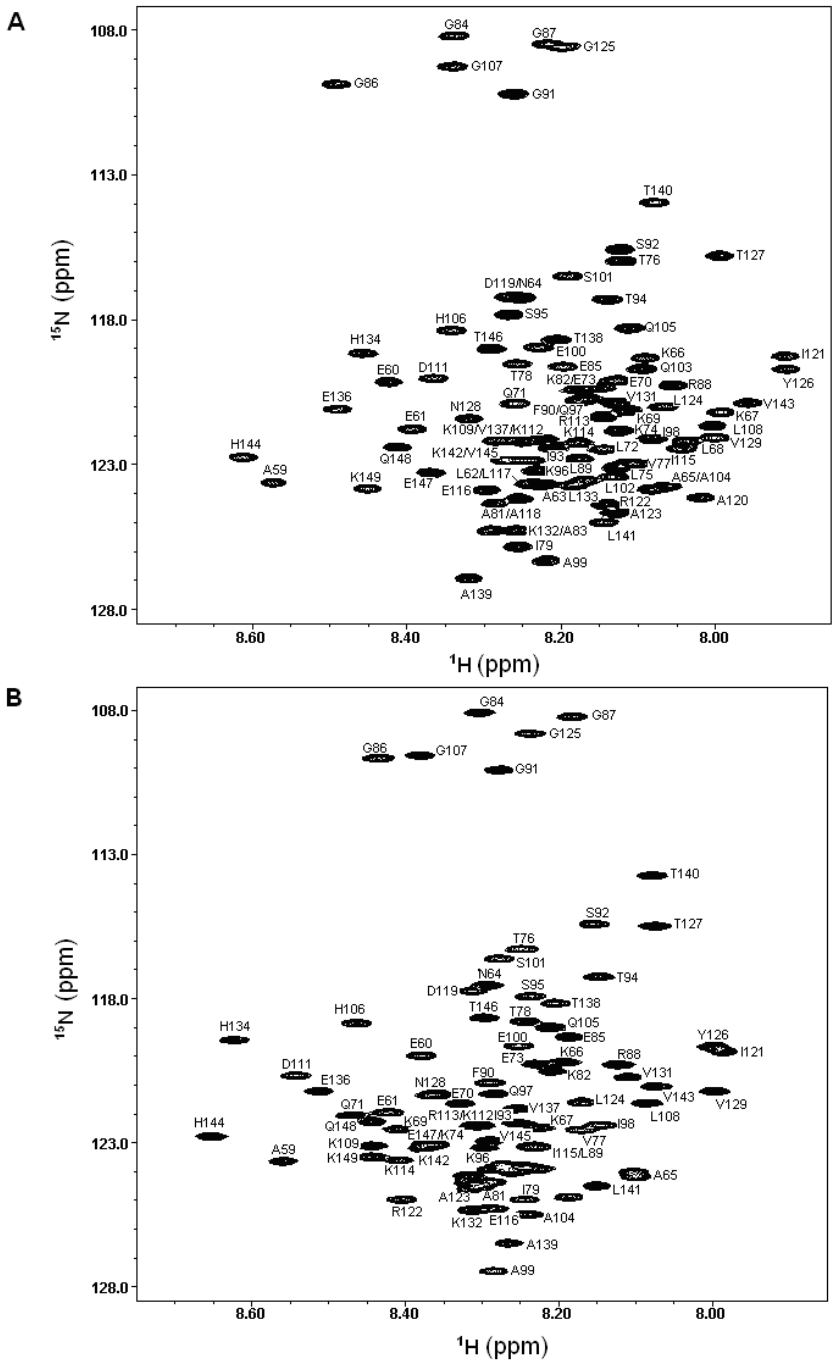


Figure 2-3. (A) The ^{15}N-^{1}H HSQC spectrum of the pH 2.0 unfolded state of CTL9. (B) The ^{15}N-^{1}H HSQC spectrum of the pH 2.5, 7.6 M urea unfolded state of CTL9. Peaks are labeled in both spectra. Spectra were collected at 25 ^{o}C.

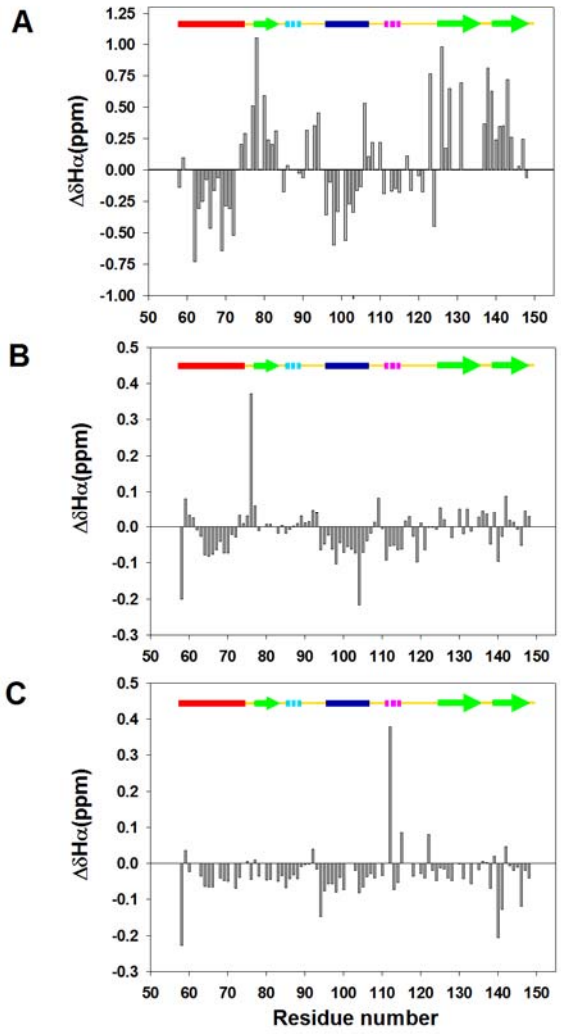


Figure 2-4. Plots of the deviations of the measured Cα ^{1}H chemical shifts and random coil chemical shifts. Data plotted as observed-random coil. Random coil values in acidic (pH 2.3) 8M urea were used (*138*) together with sequence specific corrections (*137*). (A) Deviations for the native state (pH 3.8); (B) Deviations for the pH 2.0 unfolded state. (C) Deviations for the pH 2.5, 7.6 M urea unfolded state. A schematic representation of the elements of secondary structure of the native state of CTL9 is shown at the top of the figure (arrows represent β-strands, filled cylinders represent α-helices, dashed cylinders represent 3_{10} helices and single lines represent loop regions).

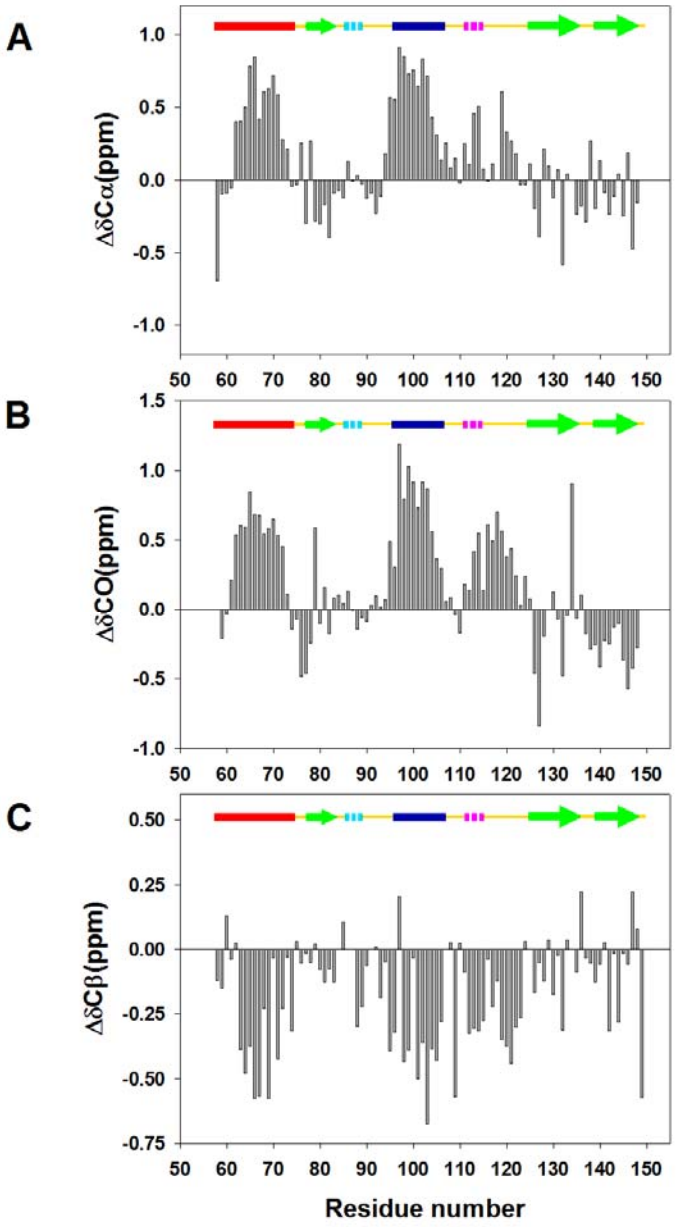


Figure 2-5. Plots of the deviations of the measured ^{13}C chemical shifts of the pH 2.0 unfolded state of CTL9 from random coil chemical values. Data plotted as observed-random coil. (A) $^{13}C\alpha$ chemical shifts deviations, (B) Carbonyl ^{13}C chemical shifts deviations, (C). $^{13}C_\beta$ chemical shifts deviations. Sequence dependent corrections were made (*137*).

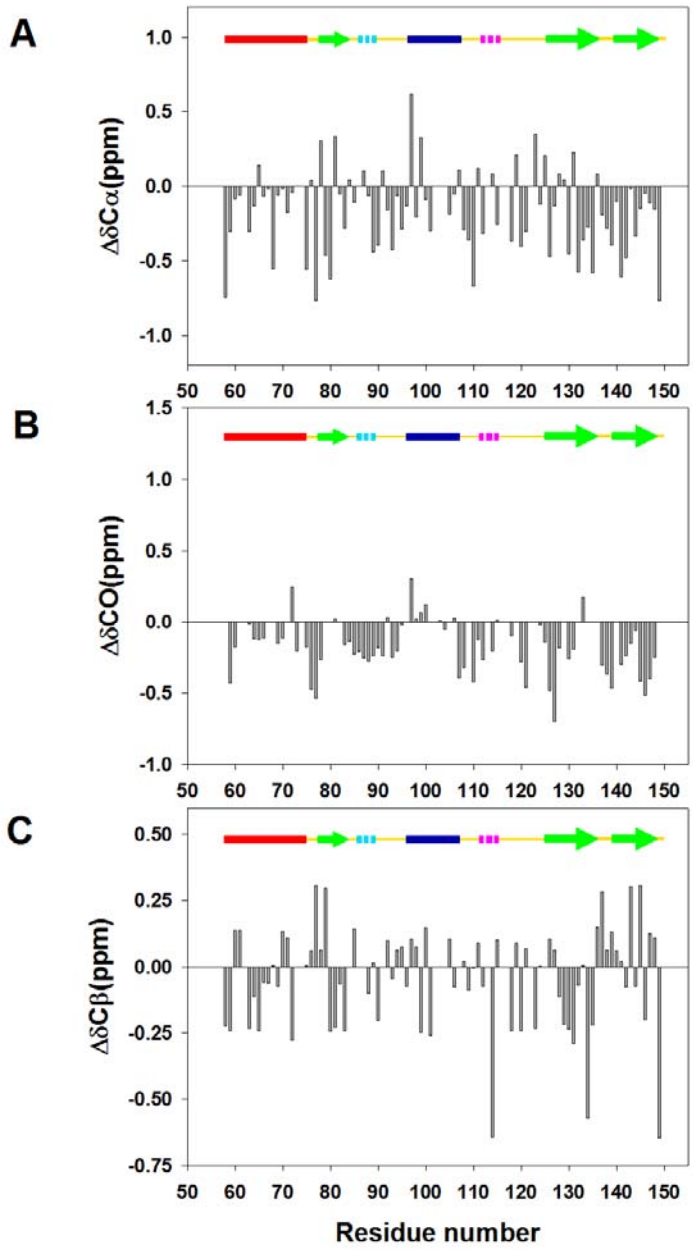


Figure 2-6. Plots of the deviations of the measured ^{13}C chemical shifts of the pH 2.5, 7.6 M urea unfolded state of CTL9 from random coil chemical values. Data plotted as observed-random coil. (A) $^{13}C\alpha$ chemical shifts deviations, (B) Carbonyl ^{13}C chemical shifts deviations, (C) $^{13}C_\beta$ chemical shifts deviations. Random coil values in acidic (pH 2.3) 8M urea were used (*138*) and sequence dependent corrections were made (*137*).

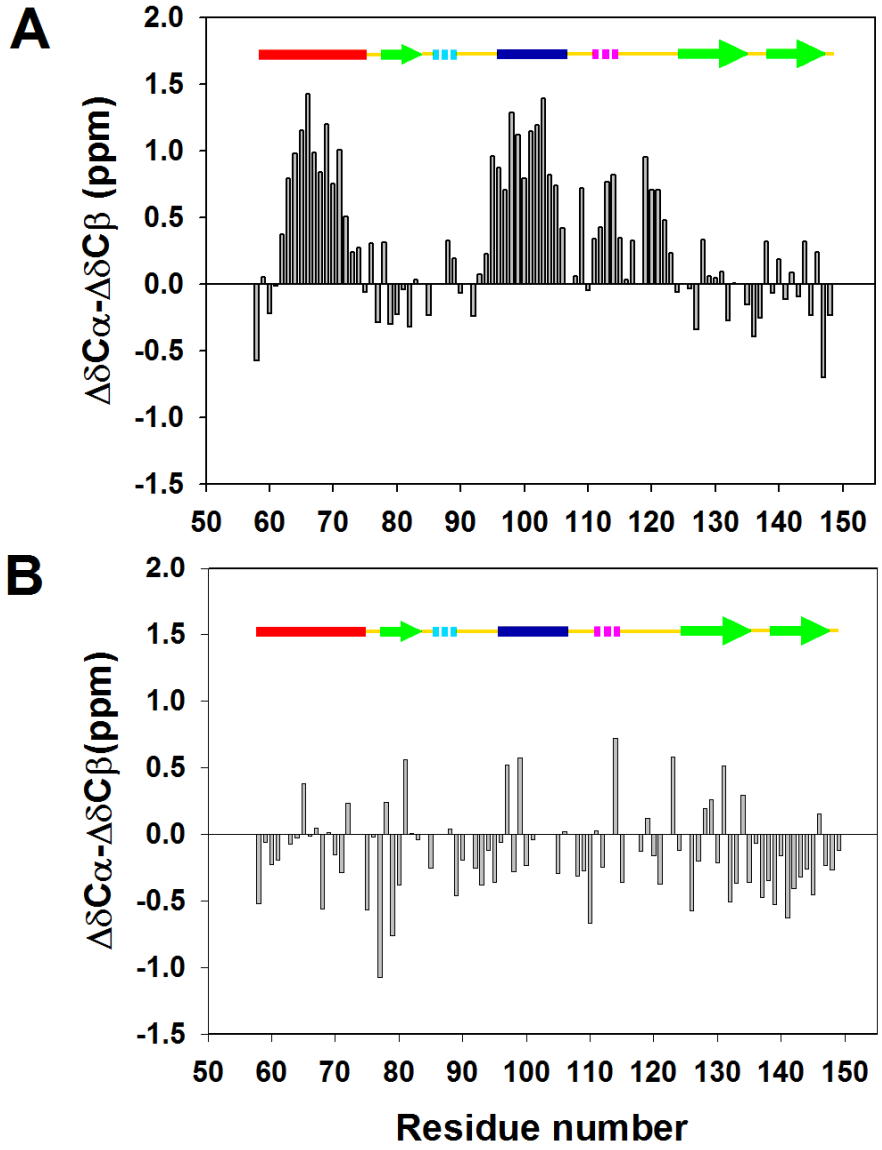


Figure 2-7. A plot of the difference of the secondary shifts of the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts for (A) the pH 2.0 unfolded state of CTL9 and (B) the pH 2.5, 7.6 M urea unfolded state. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of the figure.

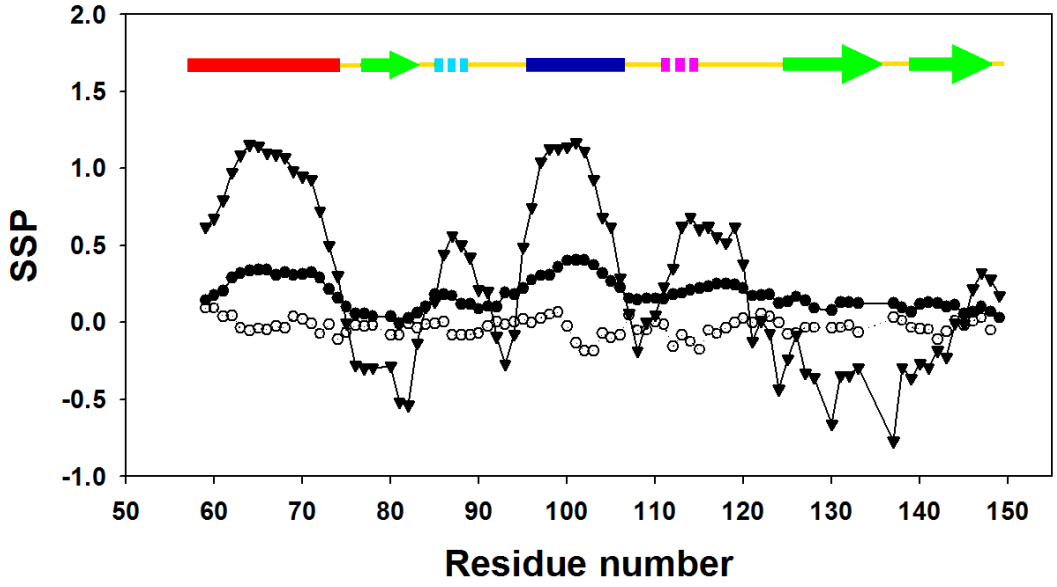


Figure 2-8. SSP analysis of the pH 2.0 (●) and the pH 2.5, 7.6 M urea (○) unfolded state of CTL9 conducted using the method of Forman-Kay and coworkers (*139*). The calculation for the pH 3.8 native state is included for comparison (▼). Positive values represent a propensity to populate the helical region of the φ, ψ map while negative values indicate a preference for the β-sheets region. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of the figure. $^{13}C_{\alpha,}$ $^{13}C_{\beta}$ and C_{α} ^{1}H chemical shifts were used in the analysis.

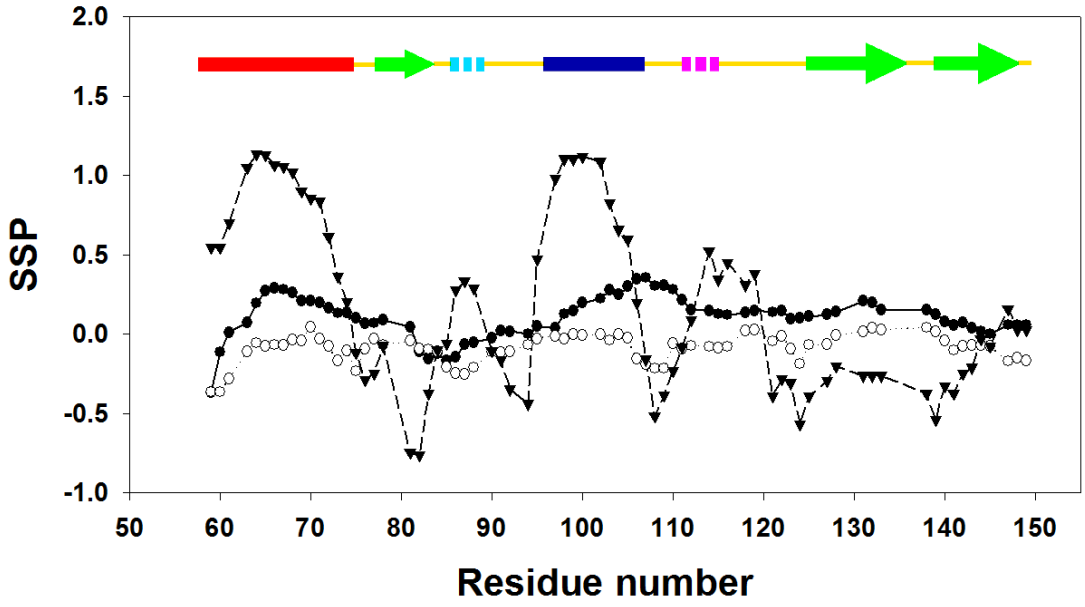


Figure 2-9. SSP analysis of the pH 2.0 unfolded state (●), the pH 2.5, 7.6 M urea unfolded state (○) and pH 3.8 native state of CTL9 (▼) calculated using only $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts.

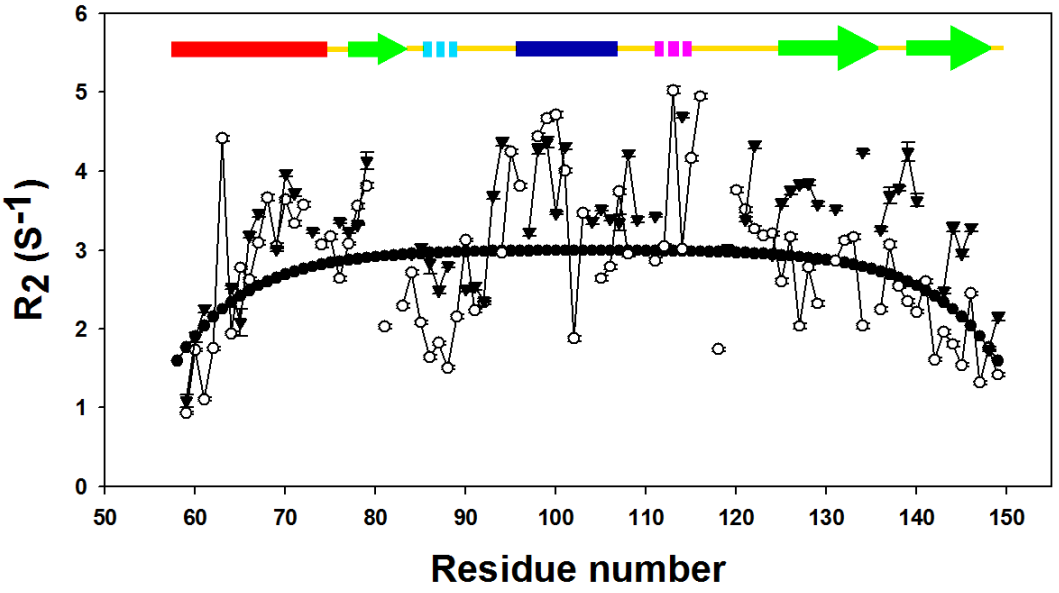


Figure 2-10. A plot of ^{15}N R_2 rates for the pH 2.0 unfolded state (○) and the pH 2.5, 7.6 M urea unfolded state (▼). The solid line is the best fit to the phenomenological model of Schwalbe and coworkers (*91*).

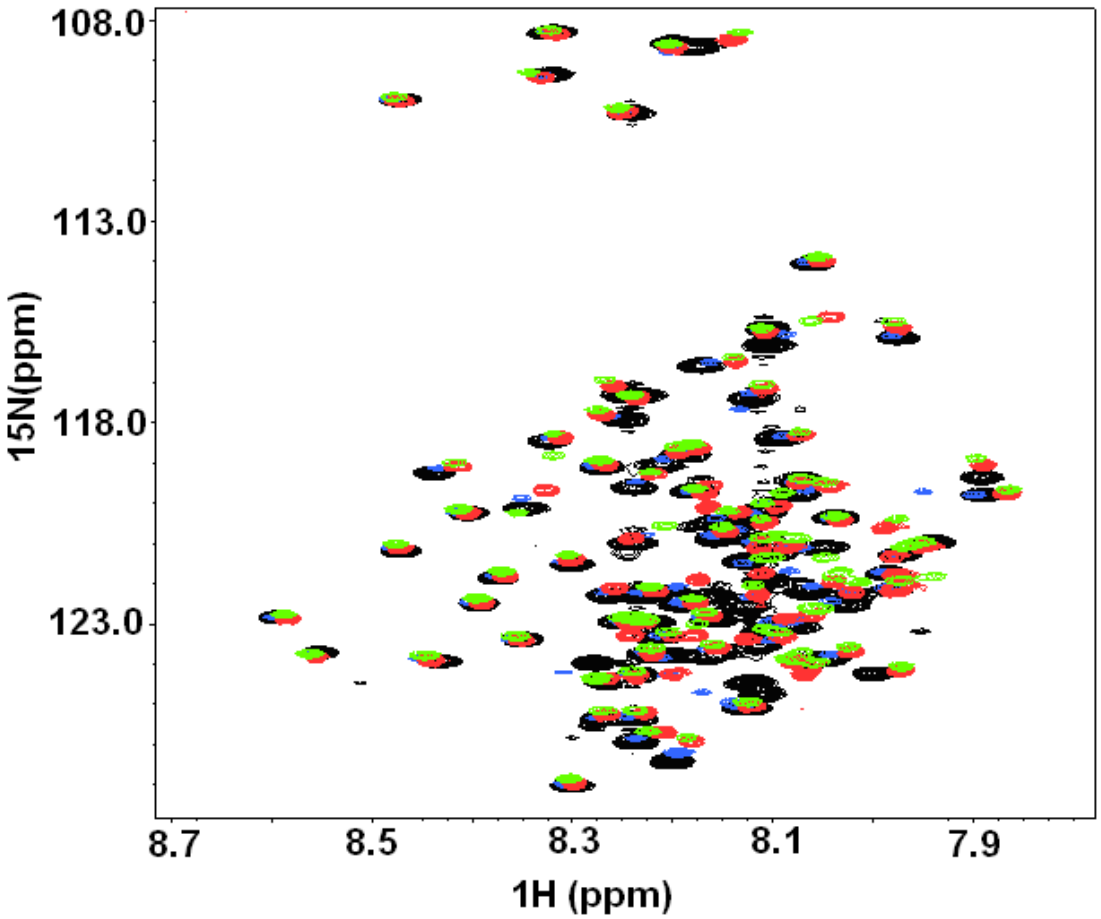


Figure 2-11. The ^{15}N-^{1}H HSQC spectra of wild-type CTL9 and its mutants. (●) Wild-type CTL9, (●) K74C CTL9, (●) K109C CTL9, (●) D119C CTL9.

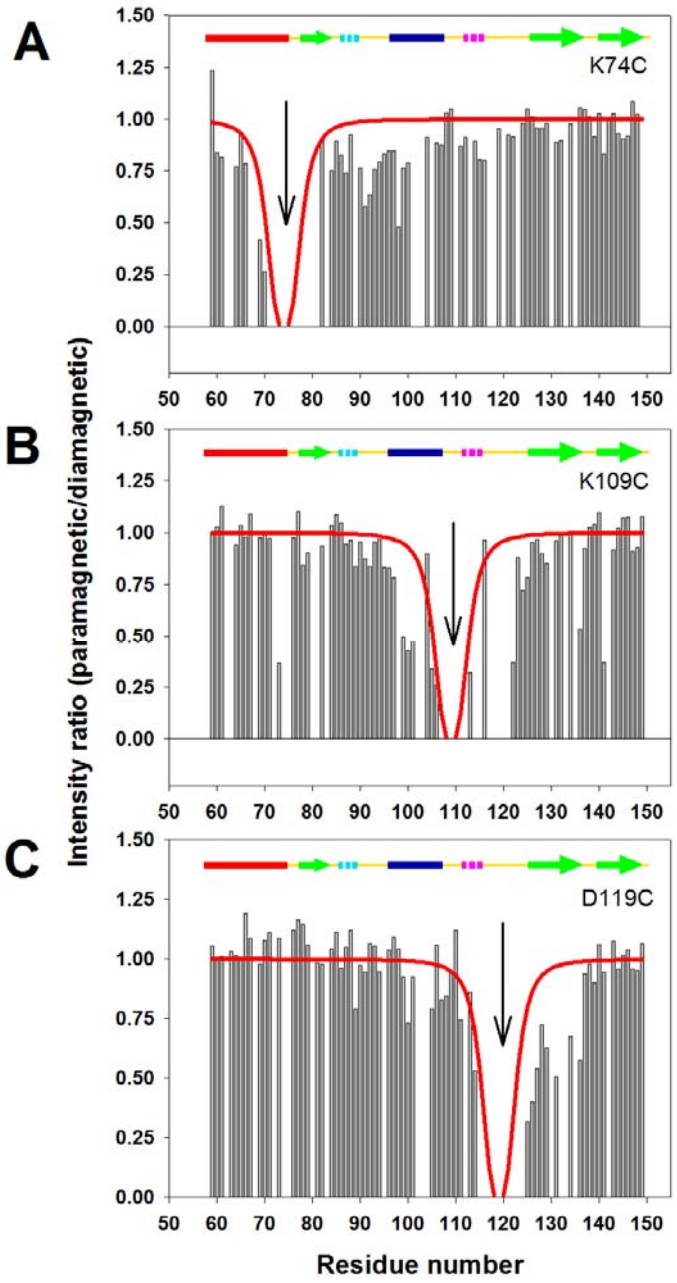


Figure 2-12. Paramagnetic enhancement to nuclear spin relaxation for the 8M urea unfolded state of CTL9 at pH 2.0. The histograms display the intensity ratio (I_{para}/I_{dia}) of the ^{1}H-^{15}N crosspeaks in the HSQC spectrum of (A) K74C, (B) K109C, (C) D119C. Continuous red lines indicate the broadening simulated using an idealized model for a random coil polypeptide chain.

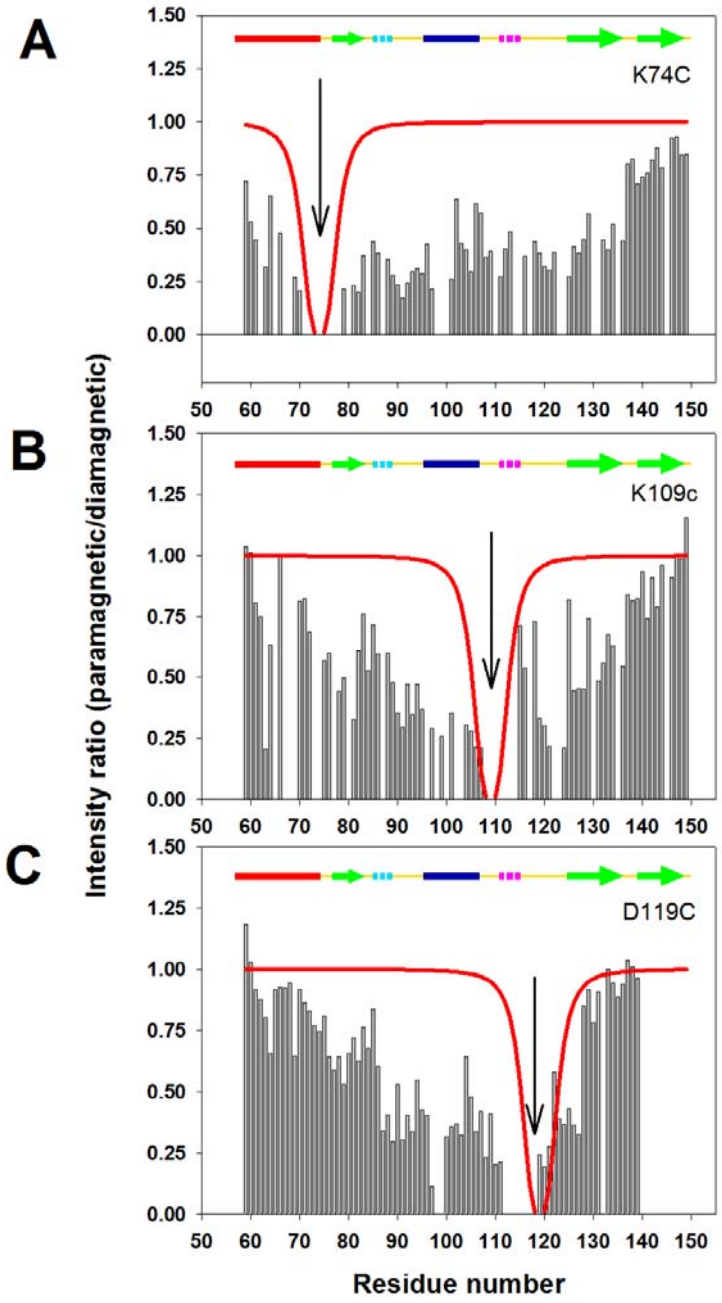


Figure 2-13. Paramagnetic enhancement to nuclear spin relaxation for the pH 2.0 unfolded state of CTL9. The histograms display the intensity ratio (I_{para}/I_{dia}) of the ^{1}H-^{15}N crosspeaks in the HSQC spectrum of (A) K74C, (B) K109C, (C) D119C. Continuous red lines indicate the broadening simulated using an idealized model for a random coil polypeptide chain.

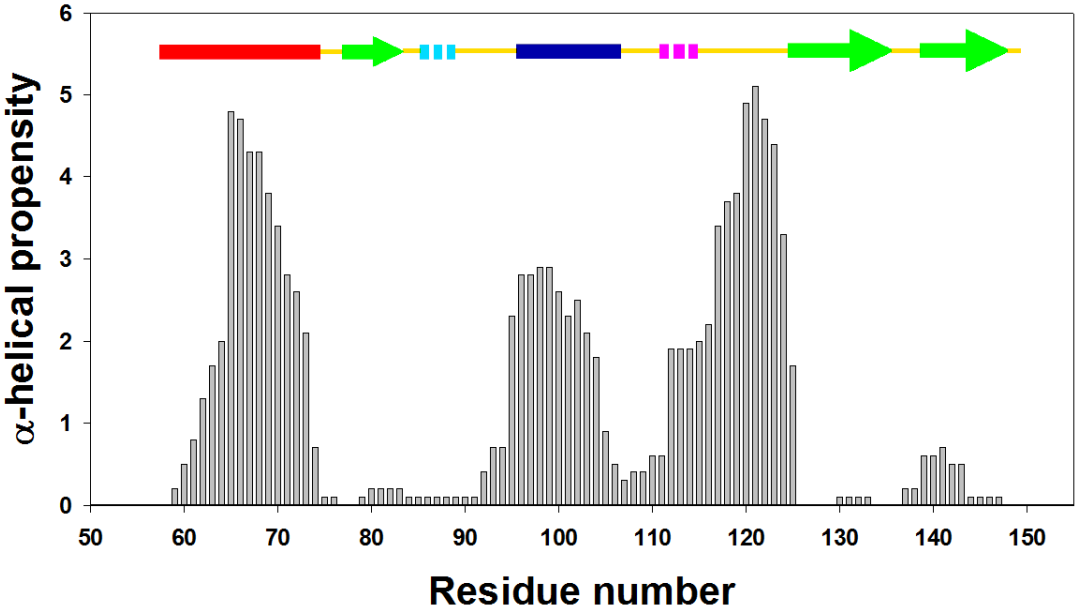


Figure 2-14. AGADIR analysis of the primary sequence of CTL9 at pH 2.0, 298 ^{o}K.

3. The unfolded state of the C-terminal domain of the ribosomal protein L9 contains both native and non-native structure.

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 as well as ^{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 which 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.

Note: The data presented here has been published (Shan, B., Eliezer, D., and Raleigh, D.P. *Biochemistry*, 2009, **48**, 4707-4719). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Prof. Raleigh and Prof. Eliezer.

3.1. Introduction

The last 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 non-random residual structure (*51, 54, 91, 110, 113-115*). 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 (*51, 53*). 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), 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 (*120-123*). High concentrations of denaturants generally abolish residual secondary structure and lead to a more expanded unfolded state. In contrast, native-like or non-native structures have been found in a variety of unfolded proteins under less harsh conditions (*78-79, 147*). Obviously, the unfolded state populated under physiological conditions is the most interesting 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 (*54, 141, 148*). 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 (*88-89, 147, 149-150*). The picture which emerged from these studies is one of an unfolded state which 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 which 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 (*54, 113, 151*). 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 (*50*). 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. The structure and thermodynamic properties of CTL9 have been well studied (*32, 103*). The protein contains three histidines, making the equilibrium unfolding strongly pH dependent (Figure 3-1), and the domain undergoes reversible unfolding as the pH changes between 7 and 2 (*32, 129*). 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 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, 8M urea) the denatured ensemble is expanded with a radius of hydration of 33.6 Å (*79*). The CD spectrum indicates that CTL9 contains significant helical structure at pH 3.8 but obviously could not assign it to one specific conformational state or localize it to specific regions of the chain.

3.2. Materials and Methods

3.2.1. Protein expression and purification

^{15}N-labelled CTL9 and ^{15}N, ^{13}C-labelled CTL9 were overexpressed in M9 minimal medium and purified as previously described (*78*). 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 per liter of media.

3.2.2. NMR experiments

Protein samples for NMR experiments were dissolved in 90% H_2O/10% D2O 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 ^{o}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 (*152*) was acquired using 1024 x 256 complex points with eight scans per increment. The spectrum 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-labelled protein sample: HNCO, HNCACO, HNCACB and CBCACONH. Typical spectrum widths were 8389.3 Hz for the ^{1}H dimension and 2270.7 for the ^{15}N dimension. Constant time HNCO (*153-154*) and constant time HNCACO (*154-155*) experiments were acquired with 1024 x 80 x 144 complex points. The ^{13}CO dimension has a spectrum width of 2817.7 Hz and centered at 176 ppm. Constant time HNCACB (*156-157*) and constant time CBCACONH (*157-158*) experiments comprised 1024 x 80 x 160 complex points. The ^{13}C dimension has a spectrum 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 3D constant time HBHACONH experiment (*157-158*) was acquired with a data matrix of 1024 (direct ^{1}H) x 80 (^{15}N) x 160 (indirect ^{1}H) complex points. The indirect ^{1}H dimension has a spectrum width of 7204.6 Hz. In all the above triple resonance experiments, watergate sequence was used to suppress water signal.

3D ^{15}N-edited HSQC-NOESY (*159*) and the ^{15}N HSQC-NOESY-HSQC (*160*) experiments were performed on a ^{15}N-labeled sample of CTL9. The ^{15}N HSQC-NOESY spectrum was acquired with spectrum widths of 9615 (HN), 2433 (^{15}N) and 9602 Hz (^{1}H) in a data matrix of 1024 x 90 x 256 complex points. The ^{15}N HSQC-NOESY-HSQC spectrum was recorded with spectrum widths of 8389 (NH), 2190 (^{15}N) and 2190 (^{15}N)

Hz and 1024 x 140 x 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 (*130*). 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. A 900 μs delay was inserted between ^{15}N 180^{o} pulses in the Carr Purcell Meiboom Gill (CPMG) pulse train. Each spectrum was collected with 4 scans using 1024 x 256 complex points. The spectrum 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 (*135*), and chemical shift assignments were made using NMRView J (*136*). 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 coworkers (*137*). The random coil values of Wishart were used to calculate the secondary chemical shifts (*87*). SSP analysis was performed using the software provided by Professor Forman-Kay free of charge on her website (http://pound.med.utoronto.ca/software.html). R2 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 (*91*).

3.3. Results

3.3.1. Sequence specific assignments 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 (*32, 78*). 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 3-2. 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 to 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 8 residues of the native state and 5 residues of the unfolded state. Overall at least partial assignments were obtained for every non-proline residue in the pH 3.8 unfolded state. All of the assignments have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/).

3.3.2. 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, so called secondary chemical shifts, are very accurate predictors of protein secondary structure (*83, 86, 161*). We used the chemical shift data reported by Wishart (*87*) 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 3-3 displays the $^{1}H_{\alpha}$ secondary shifts for CTL9. Positive $^{1}H_{\alpha}$ secondary shifts suggest α-helical structure and negative $^{1}H_{\alpha}$ secondary shifts are observed for β-sheet structure (*86*). Figure 3-3 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 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 $^{1}H_{\alpha}$ 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 $^{1}H_{\alpha}$ secondary shifts are detected from regions which 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 $^{1}H_{\alpha}$ 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 which comprise the two α-helices display contiguous negative $^{1}H_{\alpha}$ 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 to 124, also displays negative $^{1}H_{\alpha}$ 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 deviations (0.04 ppm on average). The $^{1}H_{\alpha}$ 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. $^{1}H_{\alpha}$ secondary shifts are very useful, but because of the relatively small

dispersion in ^{1}H 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 (*87, 138*). The $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{13}CO secondary shifts of the pH 3.8 native state and the pH 3.8 unfolded state are plotted in Figure 3-4 and Figure 3-5. Table 3-1 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 $^{1}H_{\alpha}$ secondary shifts. Continuous positive deviations are detected for the pH 3.8 native state (Figure 3-4A), 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 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 shows smaller amplitude $^{13}C_{\alpha}$ secondary shifts (Figure 3-5A) than the native state (Figure 3-4A), but the values are clearly significant. The largest deviations are observed in the regions which 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(*78*). The segment linking α-helix 2 and β-strand 2 (residues 107-124) also shows larger secondary $^{13}C_{\alpha}$ shifts in the pH 3.8 unfolded state (1.09 ppm) than 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 3-4B). 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 3-5B). The averages are 0.38 ppm 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 $^{1}H_{\alpha}$, $^{13}C_{\alpha}$ 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 3-4C). The average values for the three β-strands are 1.97 ppm, 1.37 ppm and 1.31 ppm, respectively, while the segment which 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.

3.3.3. 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 (*139*). Figure 3-6A 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 3-6B). 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 (*78*).

Recently, the secondary structure propensity (SSP) analysis developed by Forman Kay and coworkers has been used to analyze secondary structure propensity in natively unfolded proteins (*139*). 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 3-7 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 is 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 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 to 110 and 118 to 124 where the average deviation is 0.39. Residues 111 to 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 to 124 is 0.38 in the pH 3.8 unfolded state, and the individual values cover the range of 0.27 to 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 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 3-8. The correlation between the two pH 3.8 state 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, i.e., 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.

3.3.4. 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 3-9 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. 30 d_{NN}(i,i+1) NOEs are observed in the pH 3.8 unfolded state. These include 8 d_{NN}(i,i+1) NOEs in residues 58-74 (α-helix 1), 7 NOEs in residues 95-106 (α-helix 2), and 7 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 to 124 result from residue pairs which have inter-amide proton distances greater than 4 Å in the X-ray structure. Two of the seven result from pairs which have inter-proton distances less than 3 Å and both of these are in the regions of 3_{10} helix. This latter 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. It is clear that 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.

3.3.5. ^{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 (*90-91*). ^{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 (*90-91*). Figure 3-10 plots the measured relaxation rates and relaxation rates calculated using the phenomenological random coil model developed by Schwalbe and coworkers (*90*). As expected, the native state exhibits large systematic positive deviations from the values predicted by the phenomenological random coil model (Figure 3-10A). The deviations observed 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 3-10B). The calculated R_2 values follow the expected inverted U-shaped curve but with some noticeable deviations. Regions which 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 which 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 (equation 2) a parameter λ which has been related to the apparent persistence length of the chain. The value obtained here is 8 residues which is consistent with previous estimates (*90-91*). We note however that as rigorously defined it is not clear that the concept of persistence length is applicable to proteins.

3.4. 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 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 which 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 (*121, 123*). There are even examples of energetically significant electrostatic interactions in unfolded states (*54, 79, 82, 113, 127, 148, 162-168*). 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 (*147, 169*). 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 (*131, 170-171*).

The propensity to populate secondary structure in the pH 3.8 unfolded state is noticeably higher than that observed at pH 2.0 (*78*). 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 (*79*). 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 (*143*), an algorithm that calculates intrinsic helical propensity without consideration of the potential effects of tertiary interactions, show that it is not the case (Figure 3-11). The calculations show only a weak intrinsic helicity of the relevant CTL9 regions (3.6 % on average) that decreases only slightly (2.0 % on average) upon moving from pH 3.8 to 2.0. This very small pH-dependent decrease cannot account for the sizeable 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 according a greater degree of helix stabilization. Precedent for such an effect is provided in studies of the acid unfolded state of apomyoglobin (*131, 170-171*).

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 (*146, 168, 172*). 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, $\Delta\Delta G^{\neq}$. The dimensionless ratio $\Delta\Delta G^{\neq}/\Delta\Delta G^0$ 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. φ=1 represents an interaction which is as well developed in the transition state as in the native state while φ=0 represents an interaction which 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 which 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 (*50*). 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 which connects the two strands.

Acknowledgements

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.

Table 3-1. Average secondary shifts and SSP values in secondary structural elements of CTL9

	pH 3.8 native state					pH 3.8 unfolded state				
	$^{1}H_{\alpha}$	$^{13}C_{\alpha}$	$^{13}C_{\beta}$	^{13}CO	SSP	$^{1}H_{\alpha}$	$^{13}C_{\alpha}$	$^{13}C_{\beta}$	^{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

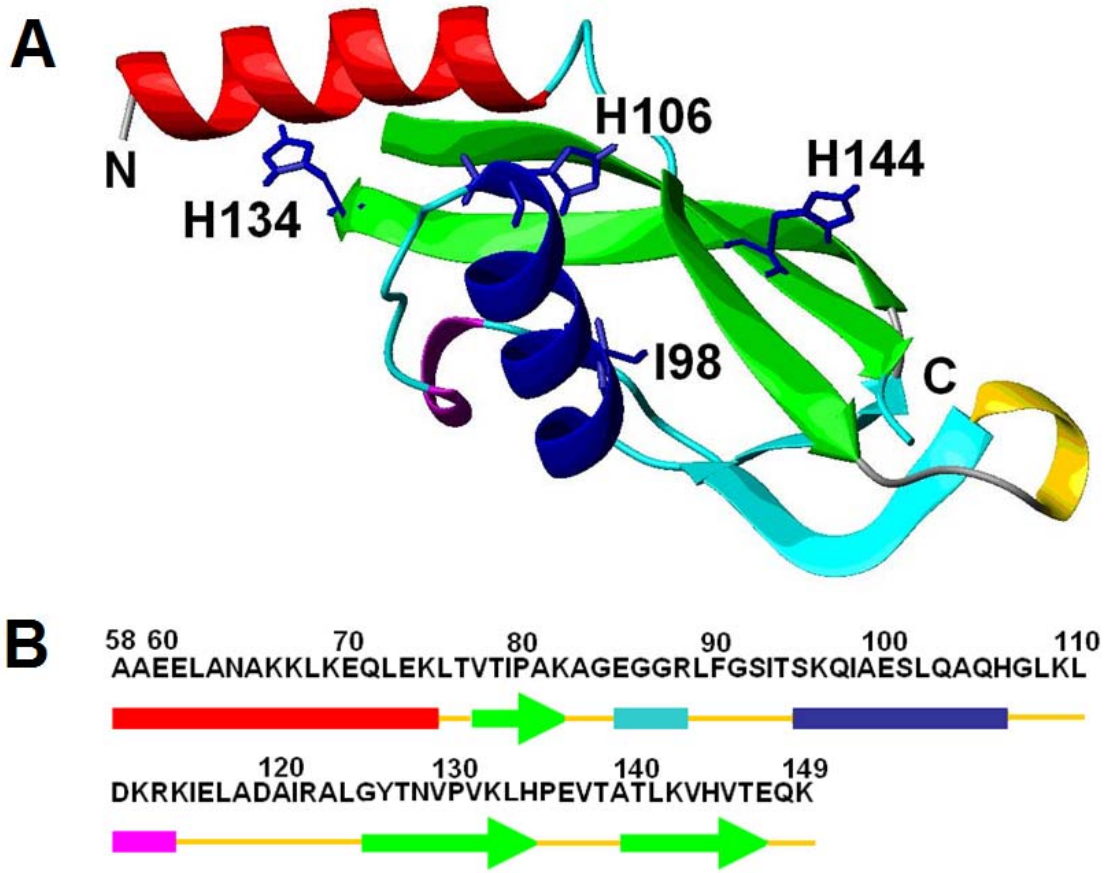


Figure 3-1. Structure of CTL9. (A) Ribbon diagram of CTL9 (residues 58-149 of protein L9). Protein data bank entry 1DIV. The N and C-termini are labeled. (B) The primary sequence of CTL9 together with a schematic representation of the secondary structural elements (arrows represent β-strands and colored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

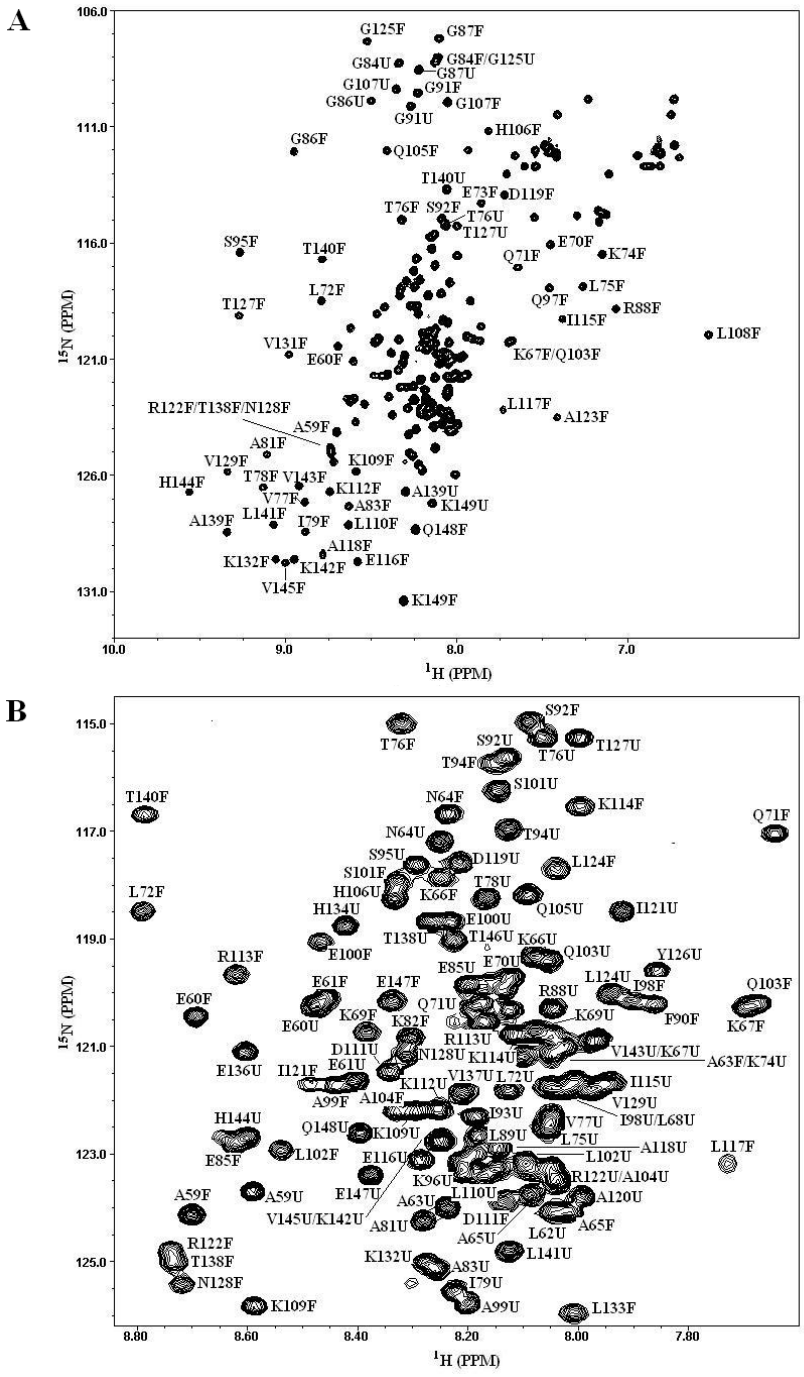


Figure 3-2. ^{15}N-HSQC spectrum (panel A) of CTL9 with assignments of peaks indicated. U denotes unfolded state resonances and F folded state peaks. The spectrum was recorded at pH 3.8 and 25 ^{o}C. Panel B is an expansion of the central region of the spectrum.

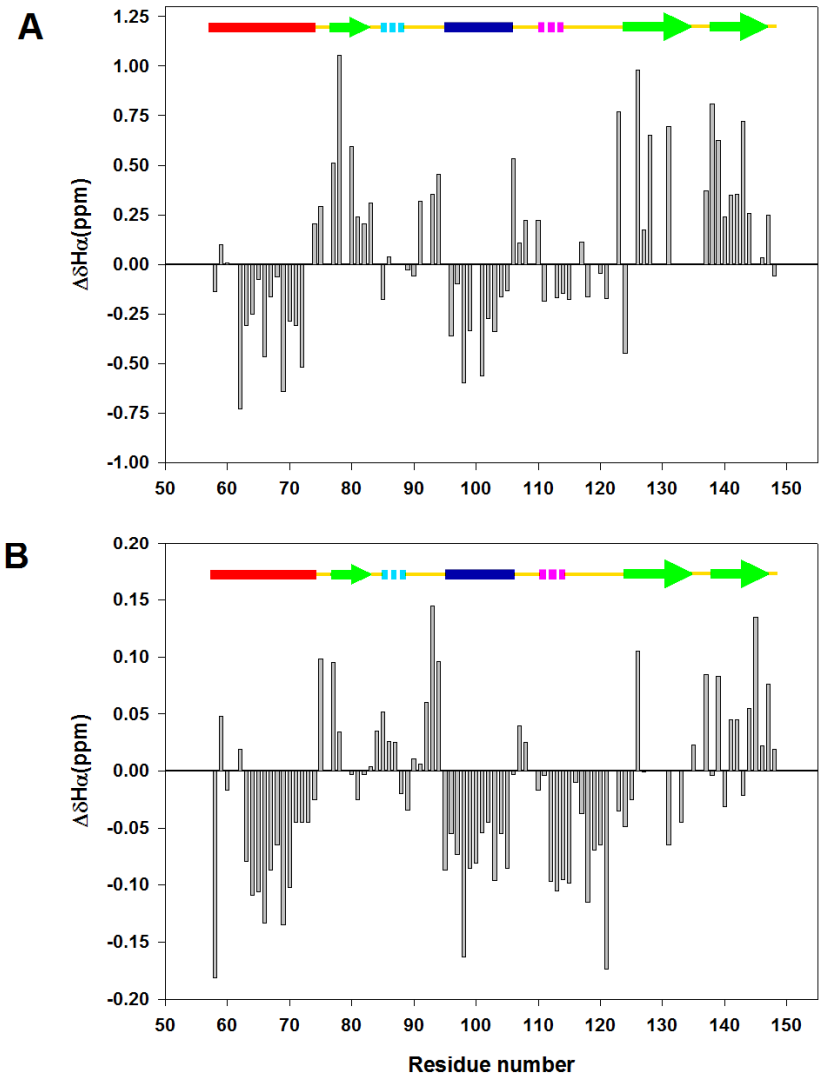


Figure 3-3. Plots of secondary ^{1}Hα shifts as a function of residue number. Random coil values from Wishart (*87*) were used together with sequence specific corrections (*137*). (A) $^{1}H_{\alpha}$ secondary shifts for the pH 3.8 native state, (B) $^{1}H_{\alpha}$ secondary shifts for the pH 3.8 unfolded 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). Values for the pH 2.0 unfolded state were taken from reference (*78*).

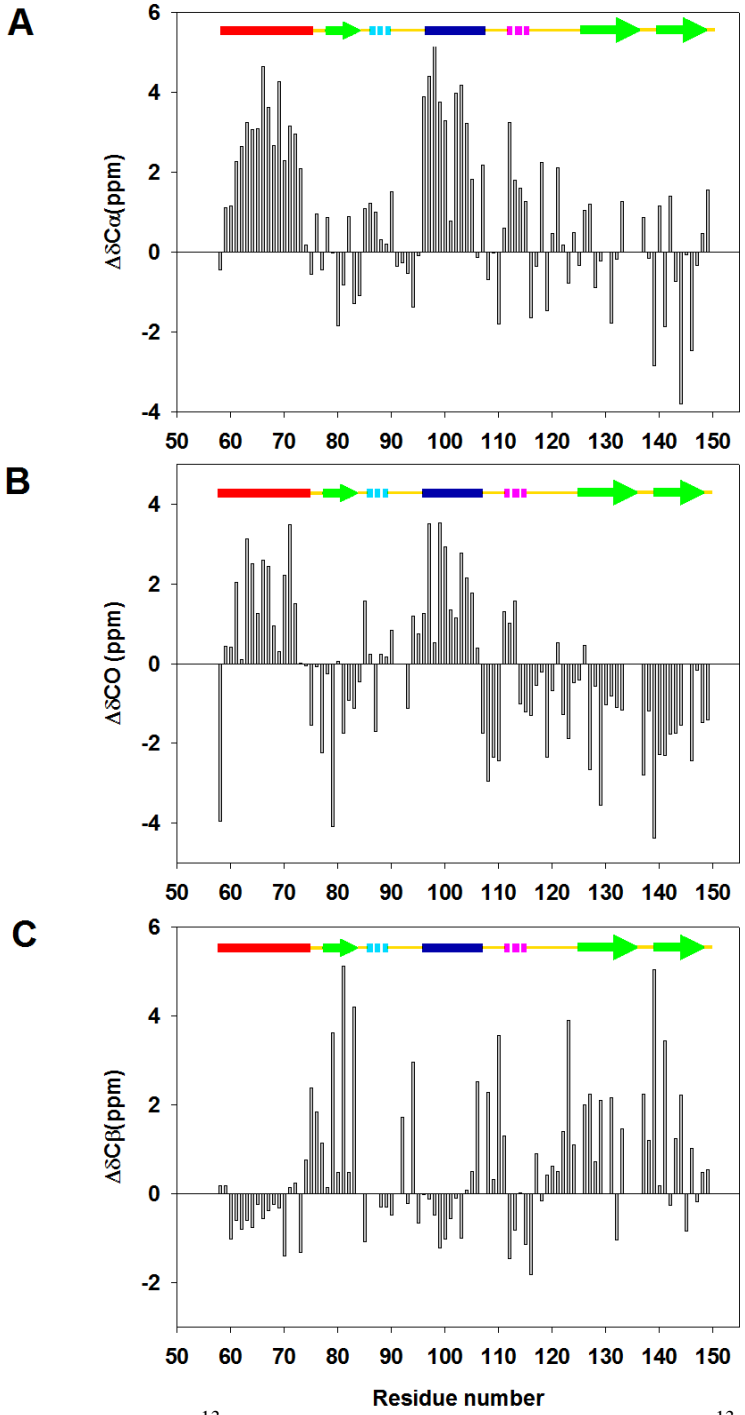


Figure 3-4. Plots of secondary ^{13}C shifts for the pH 3.8 native state. (A) $^{13}C_{\alpha}$ secondary shifts, (B) carbonyl ^{13}CO secondary shifts, (C) $^{13}C_{\beta}$ secondary shifts. Sequence-dependent corrections were made.

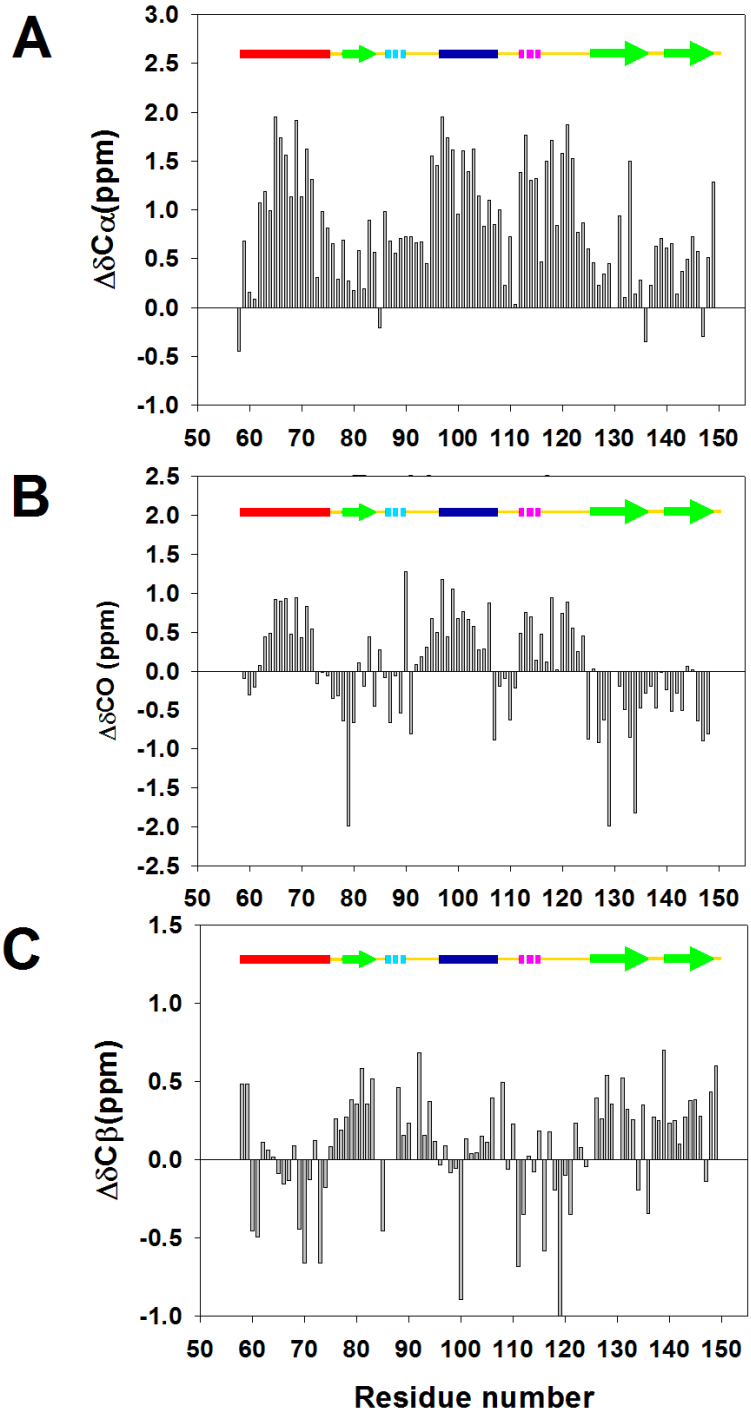


Figure 3-5. Plots of secondary ^{13}C shifts for the pH 3.8 unfolded state. (A) $^{13}C_{\alpha}$ secondary shifts, (B) carbonyl ^{13}CO secondary shifts, (C) $^{13}C_{\beta}$ secondary shifts. Sequence-dependent corrections were made for the ^{13}CO shifts.

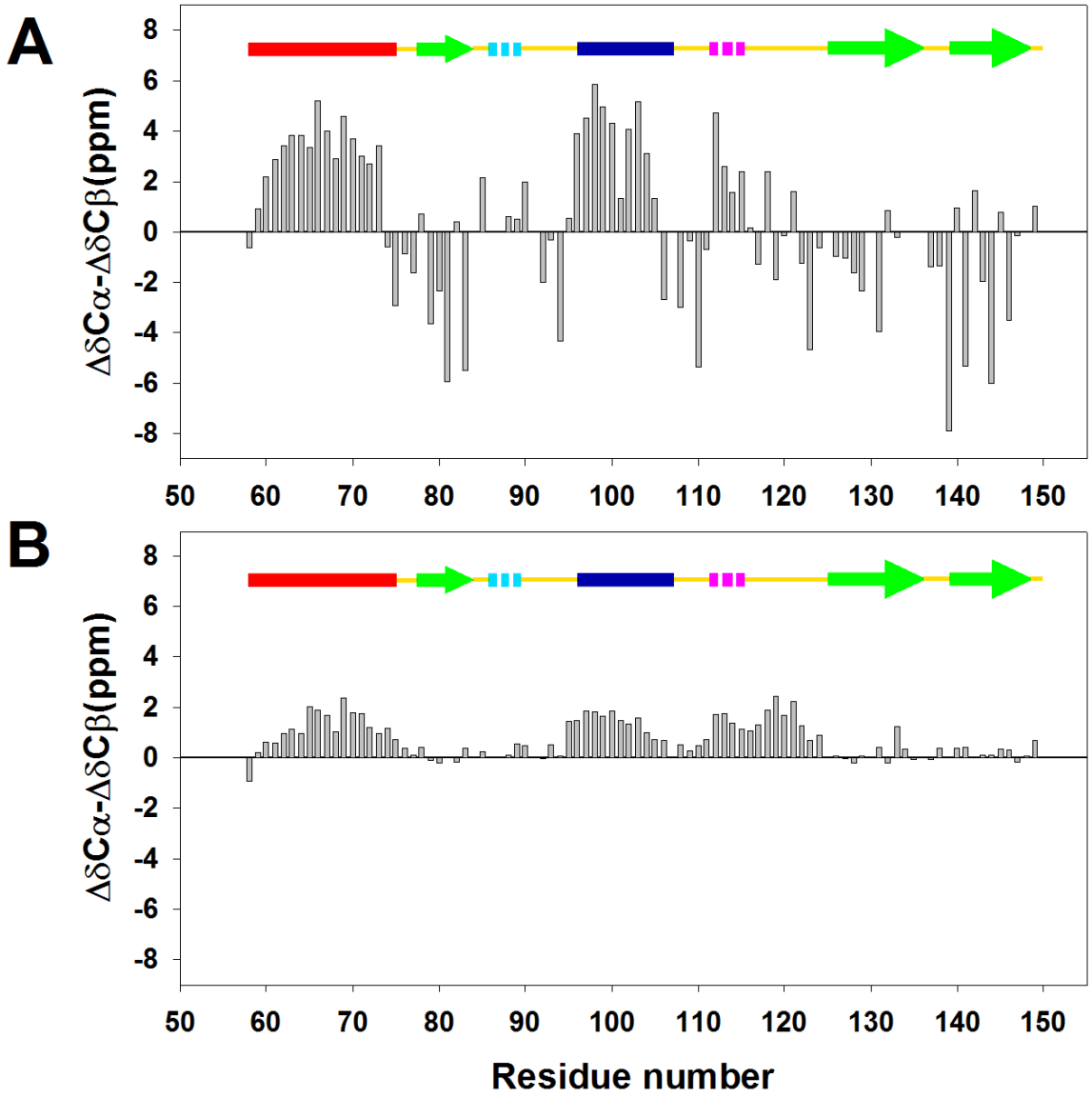


Figure 3-6. Plots of the difference in the $^{13}C_{\alpha}$ secondary shifts and the $^{13}C_{\beta}$ secondary shifts for (A) the pH 3.8 native state, (B) the pH 3.8 unfolded state. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

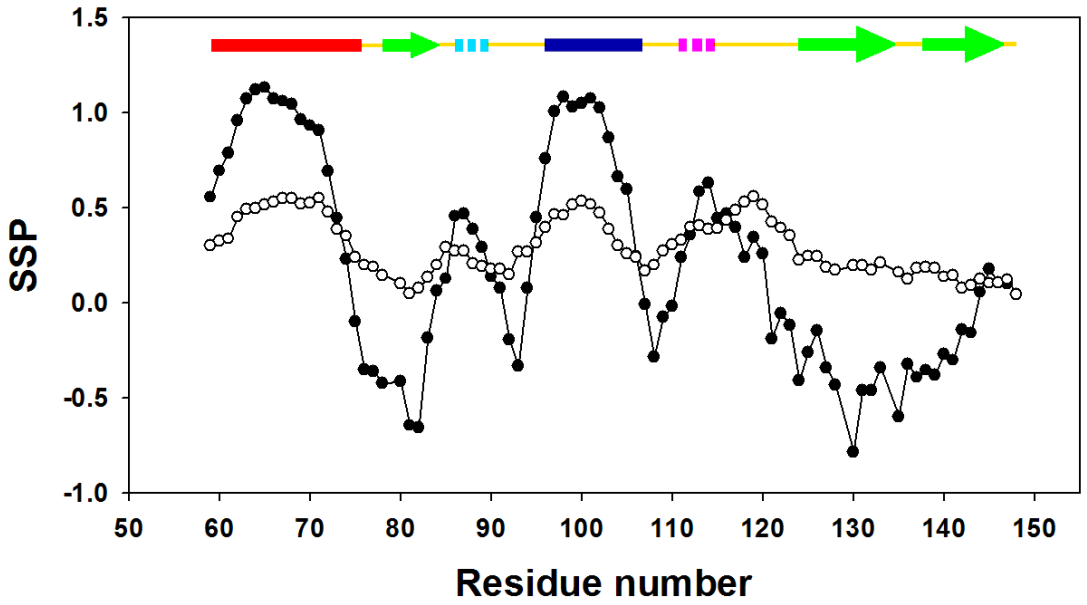


Figure 3-7. SSP analysis of the pH 3.8 native state (•) and the pH 3.8 unfolded state (o) of CTL9. Positive SSP values indicate a propensity to populate the helical region of the Ramachandran plot, while negative values indicate a preference for the β-strand region. $^{13}C_{\alpha}$, $^{13}C_{\beta}$ 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.

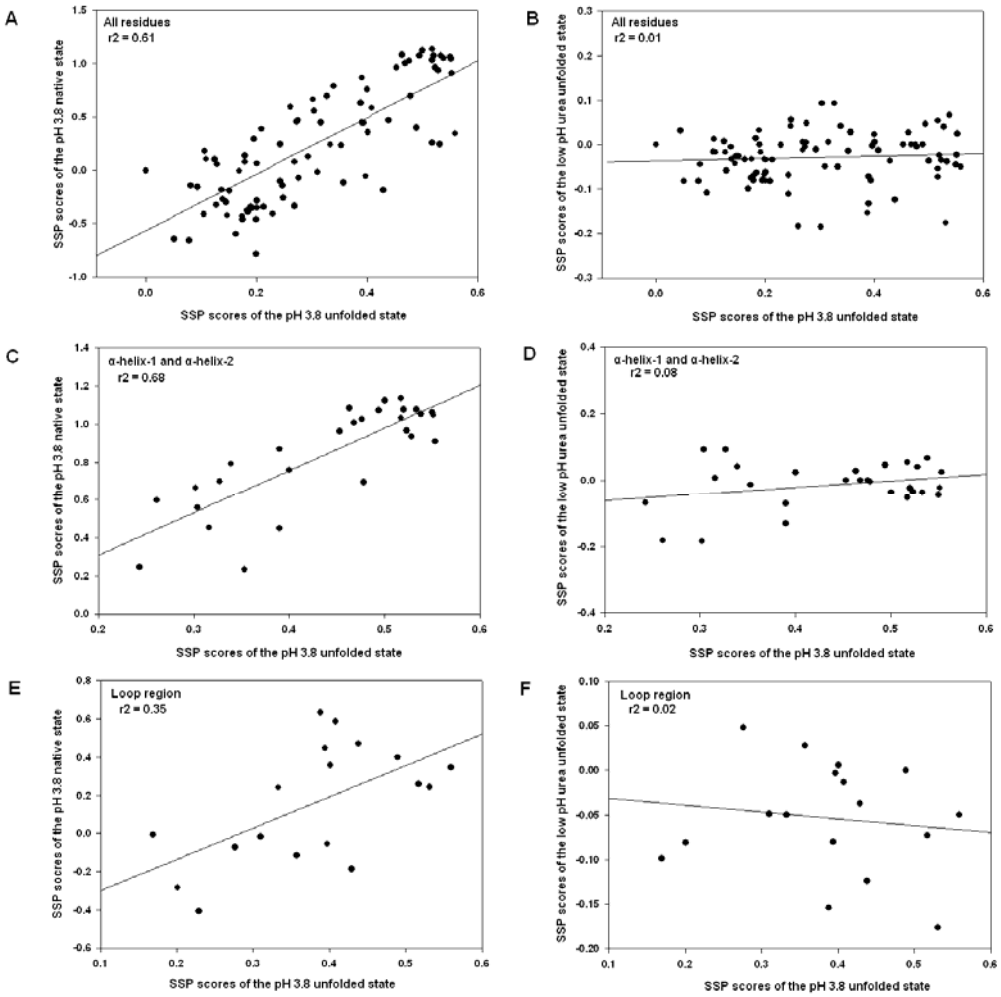


Figure 3-8. Comparison of the SSP scores of the pH 3.8 unfolded state with the SSP scores of the pH 3.8 native state and the low pH urea unfolded state. (A) Comparison of the pH 3.8 unfolded state with the pH 3.8 native state for all residues. (B) Comparison of the pH 3.8 unfolded state with the low 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.

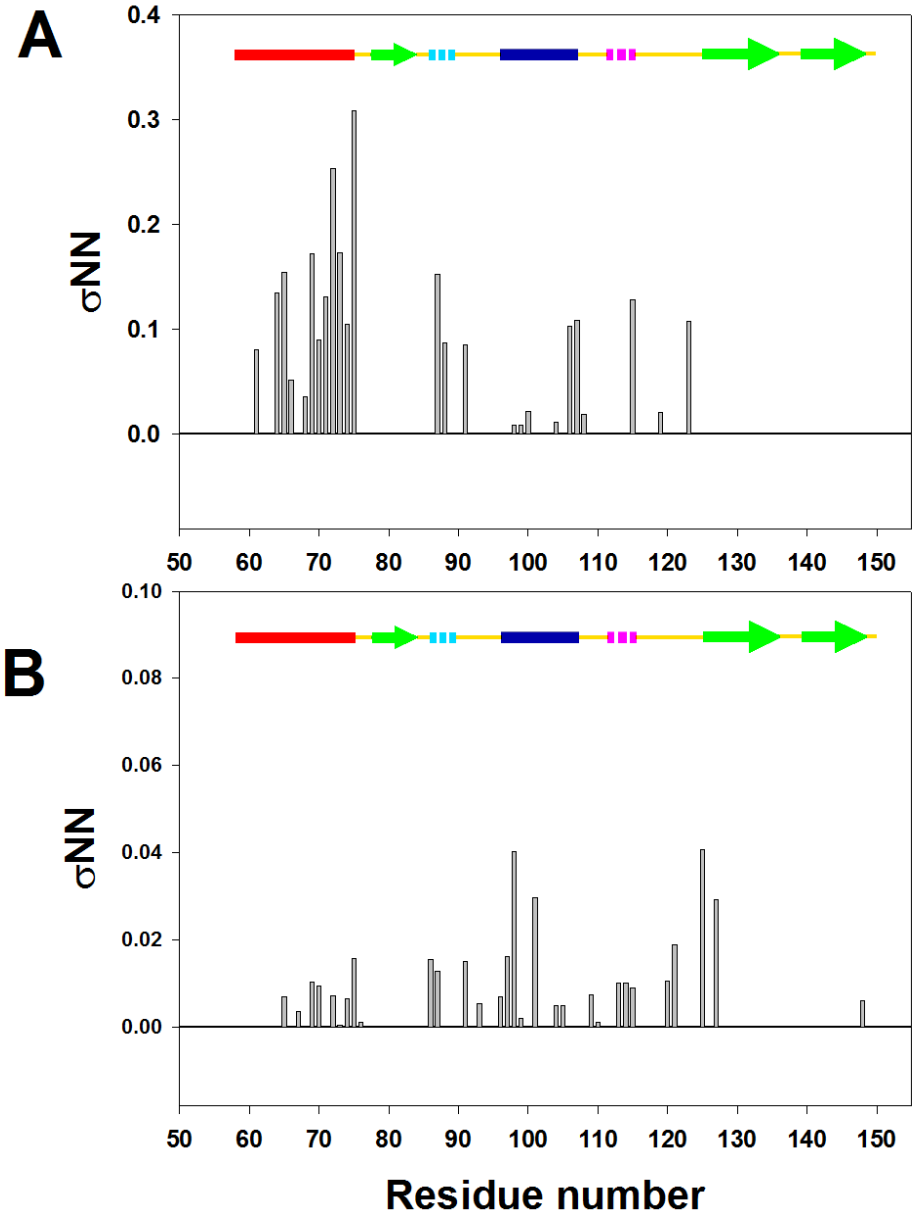


Figure 3-9. Summary of sequential d_{NN}(i,i+1) NOEs observed for CTL9 at pH 3.8. NOE peak volumes are normalized as the ratio of the volumes of the observed d_{NN}(i,i+1) NOE crosspeaks over those of the diagonal peaks. (A) the pH 3.8 native state NOEs, and (B) the pH 3.8 unfolded state NOEs. Regions where overlapping or ambiguous NOEs were observed or where no NOEs were observed are left blank.

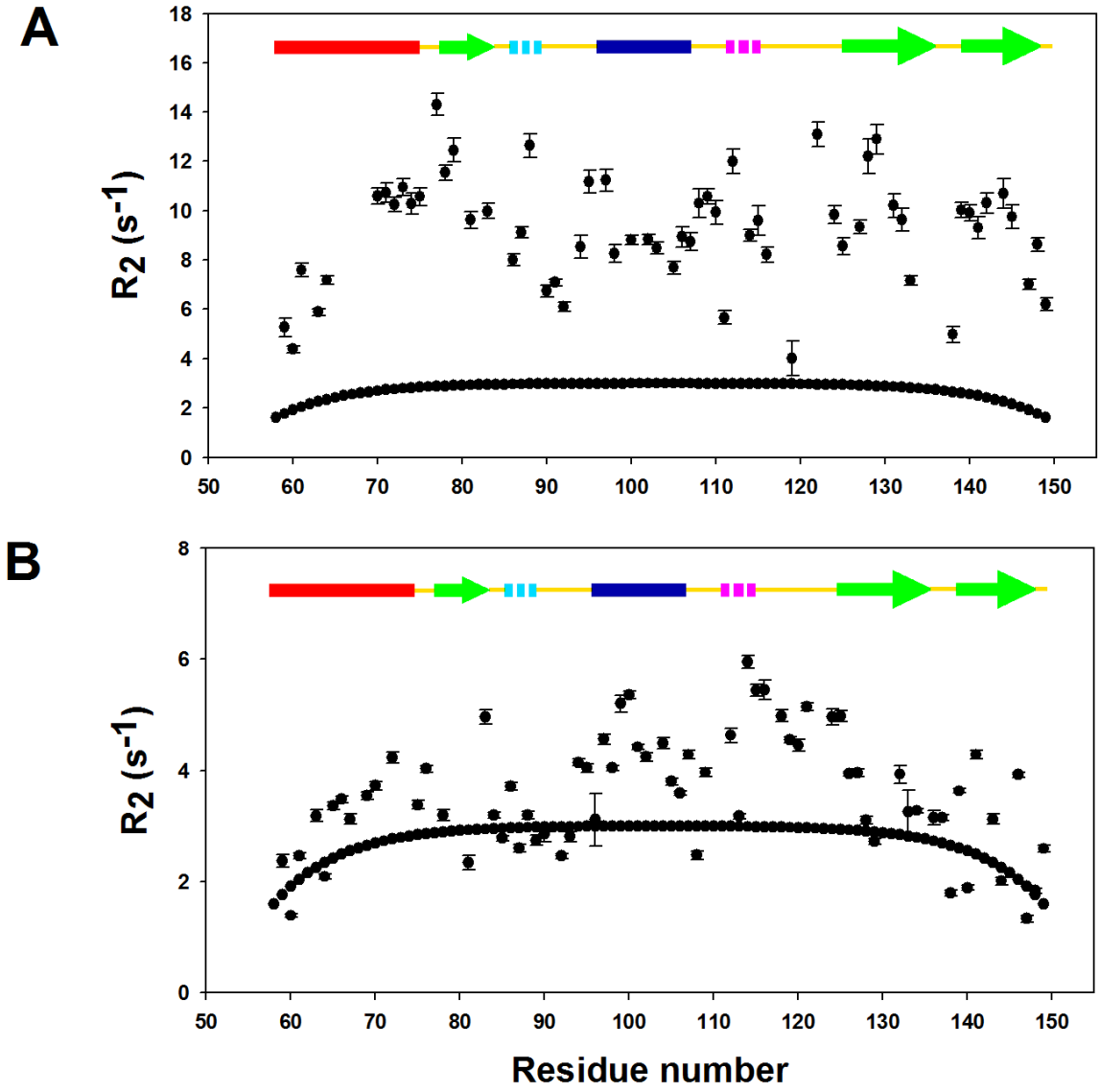


Figure 3-10. Plots of ^{15}N R_2 rates for the pH 3.8 native state (A), the pH 3.8 unfolded state (B) of CTL9. The solid line is the best fit to the phenomenological model of Schwalbe and co-workers (*91*).

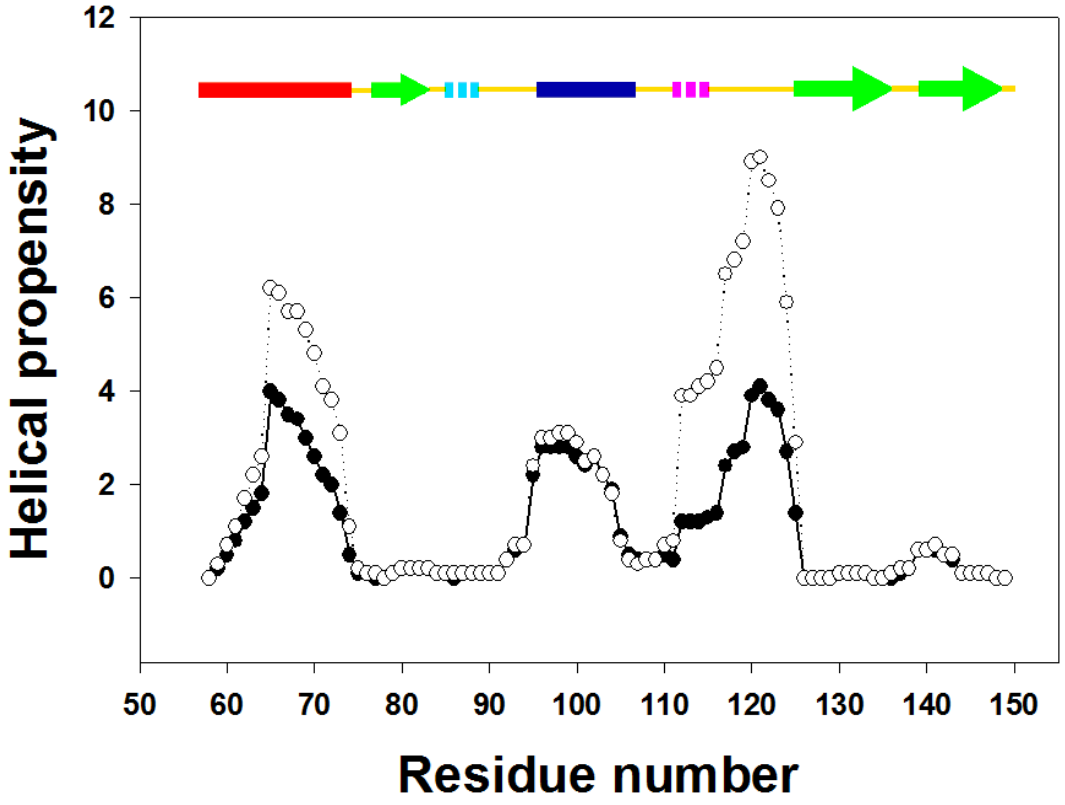


Figure 3-11. AGADIR analysis of the primary sequence of CTL9 at pH 2.0 with 20 mM salt (•) and at pH 3.8 with 120 mM salt (o). The ionic strength was chosen to match the conditions used in the respective NMR experiments.

4. The cold denaturation of the C-terminal domain of the ribosomal protein L9 is a cooperative, two-state process

4.1 Introduction

Mesophilic proteins usually undergo denaturation at temperatures both higher and lower than room temperature. While the high temperature unfolding is generally referred to as thermal denaturation, the lower temperature transition is termed cold denaturation. The phenomenon of protein cold denaturation has been observed for decades (*98*), and it has attracted increasing recent interest because the cold denaturation process has been proposed to provide access to important partially folded states (*36*) and to provide a method to thermodynamically characterize protein stability, particularly when thermal denaturation at high temperatures is not accessible (*173*). Unfortunately the fact that the mid point of cold denaturation transition is often well below the freezing point of water has prevented detailed structural and thermodynamic studies of the cold denatured state without significant additional perturbations such as addition of concentrated denaturant (*55*) , use of extreme pH (*174*), adding pressure (*175-177*) or encapsulation (*36, 178*).

It is generally believed that small single domain proteins represent extremely cooperative macroscopic systems with only two macroscopic states, i.e, the native state and the unfolded state (*98*). This view is being challenged by a number of recent studies which argue that cold denaturation is inherently highly non-cooperative and can provide experimental access to cooperative substructures and thermodynamically metastable states (*36, 179*). Such studies have used proteins encapsulated in reverse micelles and this can lead to a significant perturbation of the protein ensemble because of potential interactions between the reverse micelles and the protein. This conjecture is supported by recent findings that the cold denatured states of proteins expand at low temperatures (*108*) while the reverse micelles shrink due to the lost of water (*180*). Studies of a system under near physiological conditions are required to provide genuine insights into the cooperativity of cold denaturation.

CTL9 is a 92 residue, mixed α, β protein whose thermodynamic properties have been well described (*32, 104, 129*). The isoleucine to alanine mutation at residue 98 causes a disruption of the hydrophobic core of the protein and drastically destabilizes the protein by 4 $kcal.mol^{-1}$ (Figure 4-1). As a result, cold denaturation is observed at temperatures above 0 ^{o}C in the absence of denaturant. Both the acid induced and the urea induced denaturation of the I98A mutant follows the two-state mechanism (*129*), the cooperativity of cold denaturation are explored under native like conditions in this chapter.

4.2 Materials and Methods

4.2.1 Mutagenesis, protein expression and purification

Mutagenesis, protein expression and purification were performed as previously described for wild type CTL9 (*32, 129*). The identity of I98A CTL9 was validated using MALDI-TOF mass spectroscopy. The observed mass is 9940.9 ± 1.2 while the expected mass is 9939.5. The purity was confirmed using analytical HPLC. The yield of pure

protein was about 70-80 mg per liter of LB media and 30 mg per liter of minimal media for the I98A mutant. The expression of wild type CTL9 has a similar yield.

4.2.2 Circular dichroism spectroscopy

CD experiments were performed on a Chirascan CD spectrometer. The protein was dissolved in 10 mM sodium acetate, 150 mM NaCl buffer, pH 5.7 at a concentration of about 18 μM. Spectra were collected from 195 nm to 260 nm with a 1 nm increment and are the results of averaging 3 repeats, in a 1 mm cuvette. Thermal denaturation experiments were conducted as a function of pH between pH 4.8 and pH 8.0. Each thermal denaturation was performed by monitoring the ellipticity at 222 nm from 2 ^{o}C to 98 ^{o}C, with a 2 ^{o}C step and a heating rate of 1 ^{o}C/min, in a 1 cm cuvette. Singular value decomposition was carried out using the program implemented in the R1 software package (*181*). The pH 4.7 thermal denaturation data was fit to a quadratic equation to obtain the unfolded state signal. The pH 8.0 curve was fit to the following equation to obtain thermodynamic parameters of the native state:

$$\theta(T) = \frac{(a_n + b_n T) + (a_d + b_d T)\exp(-\Delta G_u^{o}(T)/RT)}{1 + \exp(-\Delta G_u^{o}(T)/RT)} \quad (1)$$

where a_n, b_n, a_d and b_d are parameters which define the signals of the native state (N) and denatured state (D) at a given temperature. T_m is the heat induced unfolding midpoint temperature, $\Delta H^{o}(T_m)$ is the enthalpy change at T_m, and ΔC_p^{o} is the heat capacity change between the native and denatured states. The signal expected for a fraction of folded of 0.50 was estimated by taking the average of the native state signal and the unfolded state signal.

4.2.3 Nuclear magnetic resonance experiments

^{15}N-labeled I98A CTL9 was dissolved in 90% H_2O and 10% D_2O, with 10 mM sodium acetate and 150 mM NaCl at a concentration of about 1 mM. The pH was adjusted to 5.7. ^{15}N-^{1}H correlated heteronuclear single coherence (HSQC) experiments were performed on an 800 MHz Bruker spectrometer equipped with a cryogenic probe at the New York Structural Biology from 8 ^{o}C to 30 ^{o}C with a 2 to 3 ^{o}C increment. The spectra were recorded using 1024 x 512 complex points with 8 scans per increment and spectral widths of 8802.8 Hz and 2594.7 Hz for the ^{1}H and ^{15}N dimensions, respectively. The ^{1}H dimension was centered at the water resonance and the ^{15}N offset frequency was set to 118.0 ppm. The spectra were processed using NMRpipe software (*135*).

4.2.4. Stopped-flow fluorescence

Stopped-flow fluorescence experiments were performed on an Applied Photophysics SX.18MV stopped-flow reaction analyzer equipped with an asymmetric mixing at a ratio of 10:1 (v/v). In the urea jump experiments, the proteins were denatured in 8 M urea and the refolding reaction was initiated by an 11-fold dilution into lower urea concentrations. The unfolding experiments were triggered by an 11-fold dilution into higher urea concentrations. In the pH jump experiments, the protein was initially prepared at pH 8.0 and the pH was rapidly changed to pH 5.7 the stopped flow instrument. The final concentrations of proteins were approximately 50-70 μM. The folding and unfolding reactions were detected by monitoring the changes in the fluorescence of Tyr126. An excitation wavelength of 276 nm and a 305 nm cut-off filter were used for the

measurements. Each curve was obtained by averaging 3 to 5 individual measurements and was fit using a double exponential function to obtain rate constants for two phases. The major phase corresponded to the folding/unfolding transition and was used for construction and analysis of chevron plots. The minor slow phase was believed to be due to the *cis-trans* isomerization of the three proline residues (*32*). The plot of ln k_{obs} versus urea concentration was fit to the following equation to determine the folding and unfolding rate constants in the absence of urea, $k_f(H_2O)$ and $k_u(H_2O)$:

$$\ln(k_{obs}) = \ln(k_{f(H_2O)}e^{m_f[den]/RT} + k_{u(H_2O)}e^{m_u[den]/RT}) \quad (2)$$

where m_f and m_u are constants which describe how lnk_f and lnk_u change as a function of the concentration of urea. Urea concentrations were measured by refractometry. The experiments were conducted in 50 mM NaH_2PO_4 and 100 mM NaCl in D_2O.

4.3 Results

4.3.1 Thermodynamics of cold denaturation of I98A CTL9

Similar to the wild type CTL9, the stability of I98A mutant strongly depends on pH. The protein is more stable when the histidine side-chains are deprotonated and neutral and is destabilized when the histidine side-chains are protonated at lower pH. The mid point of the pH induced unfolding occurs at about pH 5.5 at room temperature (Figure 4-2). Thermal denaturation experiments monitored by CD were conducted at various pH values. Cold denaturation is not evident at pH 8.0 in the absence of urea (*108*) but is clearly observable at lower pH values.

Thermodynamic parameters were obtained by fitting the denaturing curves to a two-state model. The $\Delta H^o(T_m)$ was determined to be 32.0 kcal.mol^{-1} at pH 6.6 and the ΔC_p^o was 1.17 kcal.mol^{-1}.K^{-1} with T_m of 49.3 ^oC (Table 4-1). In contrast, the $\Delta H^o(T_m)$ is 80.7 kcal.mol^{-1} and the ΔC_p^o was 1.07 kcal.mol^{-1}.K^{-1} for the wild type CTL9 at pH 8.0 at the T_m of 80.7 ^oC (*109*). The large decrease in the $\Delta H^o(T_m)$ reflects the fact that the truncation of the isoleucine to alanine reduces hydrophobic interactions within the core in the native state. The unfolded state of CTL9 has been shown to be relatively compact in the absence of denaturant, indicating that there might be hydrophobic clusters formed. The modest increase in the ΔC_p^o of the I98A mutant is likely due to the disruption of hydrophobic interactions in the unfolded state of the wild type caused by the mutation. Based on the Gibbs-Helmholtz equation one can derive the following expression for T_c (98):

$$T_c = \frac{T_m^{\,2}}{T_m + 2(\Delta H^o(T_m)/\Delta C_p^{\,o})} \quad (3)$$

The decrease in the $\Delta H^o(T_m)$ and the increase in the ΔC_p^o leads to an increase in the mid point of the cold denaturation. Thus it is not surprising that the cold denaturation of I98A CTL9 becomes evident at temperatures above the water freezing point. Apparent mid points of thermal (Tm) and cold denaturation (Tc) were determined based on the ellipticity expected for a fraction of folded of 0.50 (Figure 4-3). A significant increase in the Tc is observed as the pH decrease from 8.0 to 6.0 while the Tm decreases from 57.7 ^oC to 41.3 ^oC as the pH decreases (Table 4-2). The I98A mutant of CTL9 provides a system to address the cooperativity of protein cold denaturation under near native conditions.

4.3.2 Temperature dependent NMR experiments suggest two macroscopic states associated with the cold denaturation of I98A CTL9

^{15}N-^{1}H HSQC spectra were recorded for the I98A mutant of CTL9 in the temperature range from 8 ^{o}C to 30 ^{o}C, with a 2 or 3 ^{o}C change per increment, at pH 5.7 in 10 mM NaOAc and 150 mM NaCl. The spectra are displayed in Figure 4-4. The spectra collected below 12 ^{o}C are typical of an unfolded protein with limited dispersion in both the ^{15}N and the ^{1}H dimensions. Only one set of peaks are observed, indicating the cold induced unfolded state is the dominant species under such conditions. The line-widths are sharp, suggesting that the protein is monomeric which is in agreement with the hydrodynamic measurements (*108*). Two sets of peaks are observed in the spectra near room temperature, with one set corresponding to the native state and the other to the unfolded state. The peaks of the unfolded state match well with those of the cold denatured state at lower temperatures. Most resonances of the native state do not significantly shift their positions in the HSQC spectrum compared to those of the wild type protein in the native state (*182*), indicating that the I98A mutation does not perturb the native structure of the wild type CTL9 significantly. No obvious resonance broadening is detected for all the peaks, demonstrating that the two states are in slow exchange on the NMR chemical shift time scale, in agreement with our stopped-flow experiments. These studies showed that the exchange rates between the native state and the unfolded state are on the seconds time scale. The exchange rates were 6.5 S^{-1} and 1.3 S^{-1}, respectively at pD 6.6 and 15 ^{o}C in the absence of urea. The exchange rates are slightly slower at pD 8.0 in the presence of 1.4 M urea or at pH 5.7 (Table 4-3). The shape of the Chevron plot is consistent with two state folding (Figure 4-5). At intermediate temperatures between 25 ^{o}C and 12 ^{o}C , the resonances of the native state gradually attenuate and the peaks of the unfolded state increase with decreasing temperature, in an approximately synchronous manner.

4.3.3 The cold denaturation of I98A CTL9 is cooperative

To test the cooperativity of the cold induced unfolding, temperature dependent far-UV CD spectra were recorded for I98A CTL9, as shown in Figure 4-6. The spectrum at 25 ^{o}C indicates a mixture of α-helix and β-strand. At temperatures lower than 10 ^{o}C, the spectra become characteristic of an unfolded protein. The spectra display an isodichroic point at 206 nm. Singular value decomposition (SVD) was used to analyze the temperature dependent CD spectra. In an SVD analysis, the number of significant components represents the minimum number of linearly independent spectra needed to reconstruct the entire dataset. In other words, the number of significant SVD components reveals the number of detectable subensembles of conformers, i.e, the number of significantly populated thermodynamic states (*183-185*). As shown in Figure 4-7, SVD analysis for the cold denaturation of I98A CTL9 can be described by only two significant singular components. The singular vector with the second largest singular value has only 14 percent of the amplitude of the singular vector with the largest singular value. The third spectrum has a relative amplitude of only 2%. All other components are negligible. Although it is hard to describe the physical implications of the singular vectors, the appearance of the isodichroic point on the temperature dependent CD spectra as well as the SVD analysis both indicate that the cold denaturation of I98A CTL9 is a cooperative, two-state process.

4.4 Discussion

There is a strong consensus about the high cooperativity of protein folding for denaturant or thermally induced unfolding, especially for globular single domain proteins. However, studies of the cold denaturation of ubiqintin encapsulated in reverse micelles have led to debate on the cooperativity of cold denaturation. The data presented here shows that the cold unfolding of CTL9 is a cooperative two state process. Thermodynamic experiments demonstrate that the destabilizing single point mutant, the I98A mutant of CTL9, has a decreased $\Delta H^o(T_m)$ and increased ΔC_p^o and exhibits cold denaturation at temperatures above zero ^oC in the absence of denaturant under near neutral pH. Our stopped-flow experiments showed that the folding kinetics of I98A CTL9 follows a V-shaped plot and no significant deviation from linear behavior is observed (Figure 4-5). Temperature dependent NMR experiments, temperature dependent CD spectra as well as the SVD analysis all indicate that the cold denaturation is a cooperative process which features only two macroscopic states, i.e., the native state and the cold unfolded state.

Why do the results reported here differ from studies of ubiquitin encapsulated in reverse micelles? The likely explanation is that the studies in reverse micelles are complicated by temperature dependent protein micelle interactions. As reported by Flynn and coworkers, reverse micelles shrink and shed water at low temperatures at rates which are dependent on the ionic strength of the system (*186*). In addition, our PFG-NMR diffusion experiments have demonstrated that the cold denatured states of proteins expand at low temperatures (*108*). Thus apparent cold denaturation in reverse micelles is likely complicated by protein micelle interactions. It is extremely difficult to distinguish protein micelle interactions and true cold denaturation under physiologically low ionic strength (*178*). Halle and coworkers have recently monitored the hydration dynamics of ubiquitin in non-perturbing picoliter emulsion droplets using water-^{17}O spin relaxation (*180*). Ubiquitin was found to be thermodynamically stable even at -32 ^oC. Their study shows that the cold denaturation of ubiquitin encapsulated in reverse micelles is induced by the low water content in the micelles rather than by low temperature. Thus, the non-cooperativity of cold denaturation is likely an artifact of the system used in such studies.

The most important result of the current study is that we have clarified the high cooperativity of protein cold denaturation under native conditions. This is of significance given the realization that the access to cold denaturation under native conditions is potentially useful for characterizing thermodynamics and protein stability especially when denaturation by other means is not accessible (*173*).

Acknowledgements

Mr. Vadim Patsalo is greatly appreciated for performing the Singular Value Decomposition (SVD) analysis.

Table 4-1. Thermodynamic parameters for the unfolding of wild-type CTL9 and the I98A mutant determined from CD monitored thermal denaturation experiments.

	Wild type at pD 8.0	I98A CTL9, pH 8.0	I98A CTL9, pH 6.6
T_m (^oC)	80.7	58.2	32.0
$\Delta H^o(T_m)$ (kcal mol^{-1})	69.0 ± 2.4	46.1 ± 0.2	49.3 ± 0.6
ΔC_p^o (kcal mol^{-1} deg^{-1})	1.07 ± 0.08[a]	1.23 ± 0.06	1.17 ± 0.05

[a.] Date from reference (*109*). The measurements were made in 20 mM sodium phosphate and 100 mM NaCl. The quoted uncertainties are the standard error to the fit.

Table 4-2. Mid points of thermal (Tm) and cold denaturation (Tc) of I98A CTL9 at different pH values estimated based on apparent ellipticity when 50 % of the protein molecules are folded and 50% unfolded.

	I98A CTL9 pH 8.0	I98A CTL9 pH 6.6	I98A CTL9 pH 6.0
T_c (^oC)	-9.5	-0.3	16.0
T_m (^oC)	57.7	51.3	41.3

Table 4-3. Kinetic parameters for the unfolding and refolding of I98A CTL9 at various temperatures as determined by stopped-flow experiments

	Observed first order rate constant k_{obs} (s^{-1})		
	pH 5.7, no urea[a]	pD 6.6, no urea[a]	pD 8.0, 1.4 M urea[b]
25°C	1.96 ± 0.09	6.5 ± 0.01	2.0 ± 0.2
15°C	0.5 ± 0.00[c]	1.3 ± 0.01	0.5 ± 0.05

[a.] Experiments were performed in 50 mM sodium phosphate, 100 mM NaCl buffer. Standard errors to the fits are given.
[b.] Observed rate constants which were derived from fits to a double exponential, correspond to the sum of the folding and unfolding rates in 1.4 M urea. The standard error for the determined folding and unfolding rates are less than 10%.
[c.] Determined using a pH jump experiments at 12 ^{o}C.

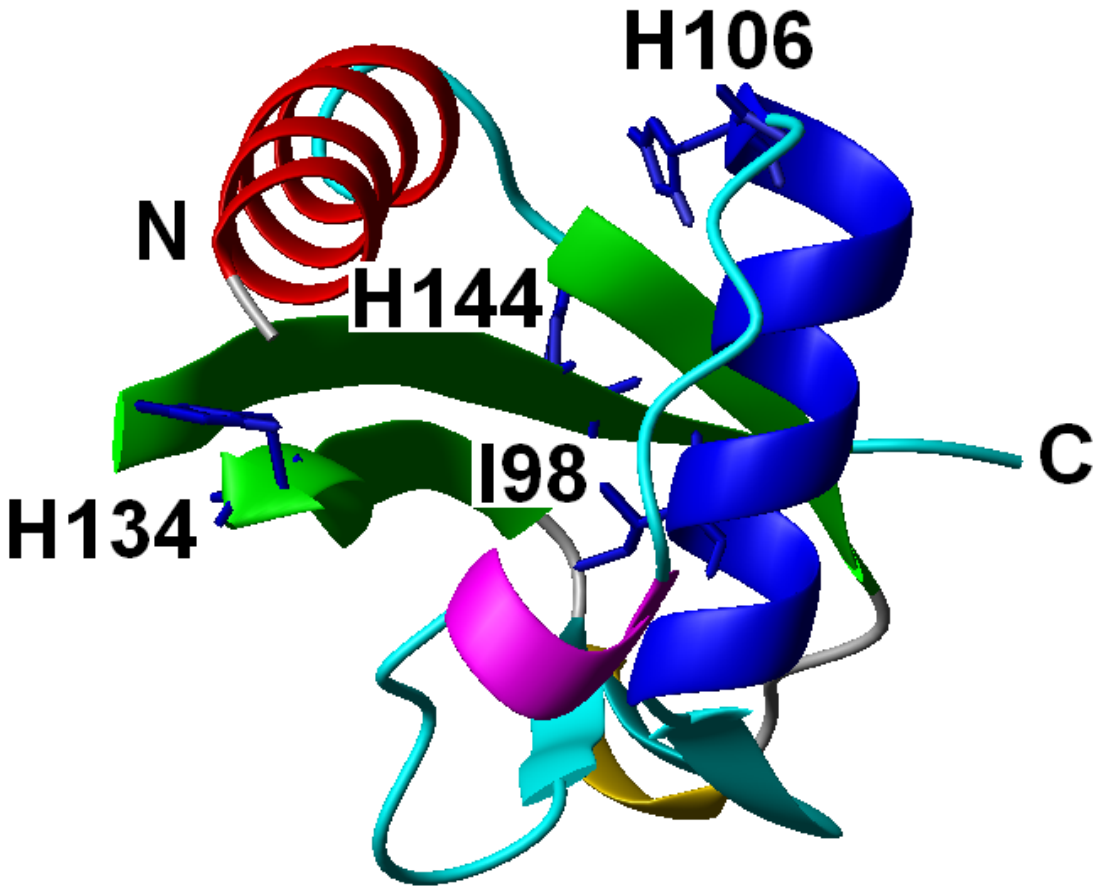


Figure 4-1. A ribbon diagrams of CTL9 made using PyMol. Side-chains of the three histidines, the hydrophobic core residue I98 and the two termini are labeled. The PDB file 1DIV was used.

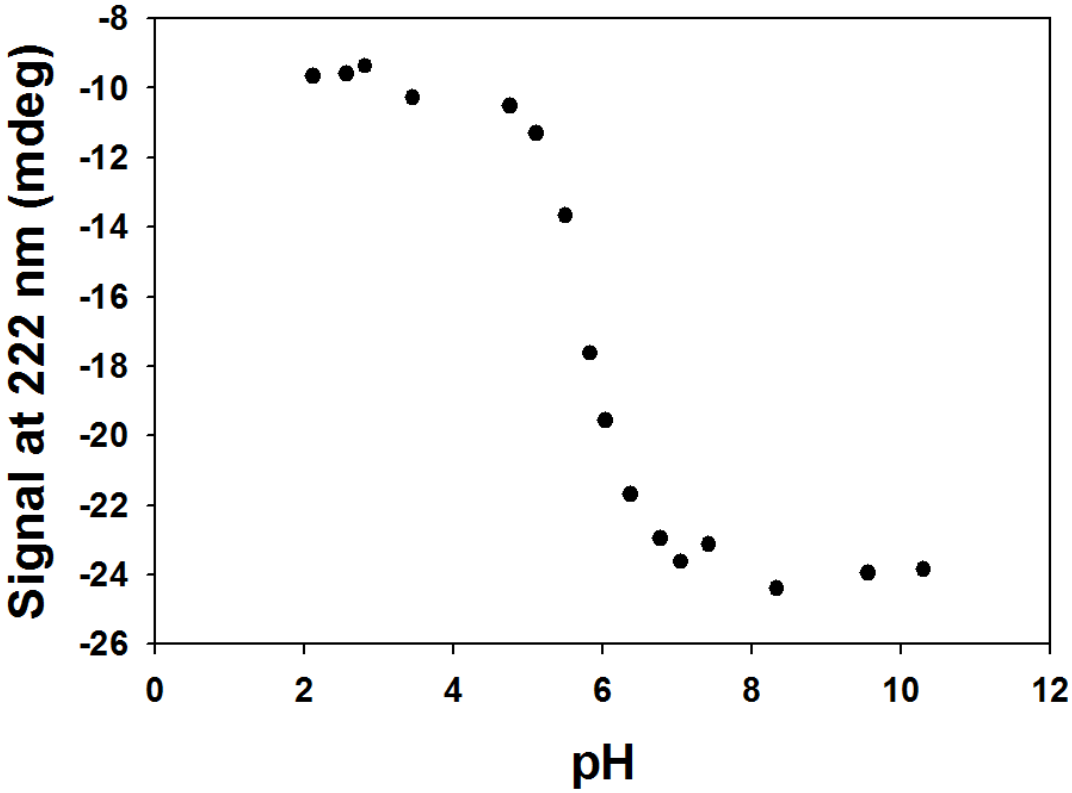


Figure 4-2. pH-induced denaturation of I98A CTL9. CD ellipticity at 222 nm was monitored in 10 mM NaOAc and 150 mM NaCl in H_2O. The pH values were adjusted by adding HCl and NaCl directly into the cuvette.

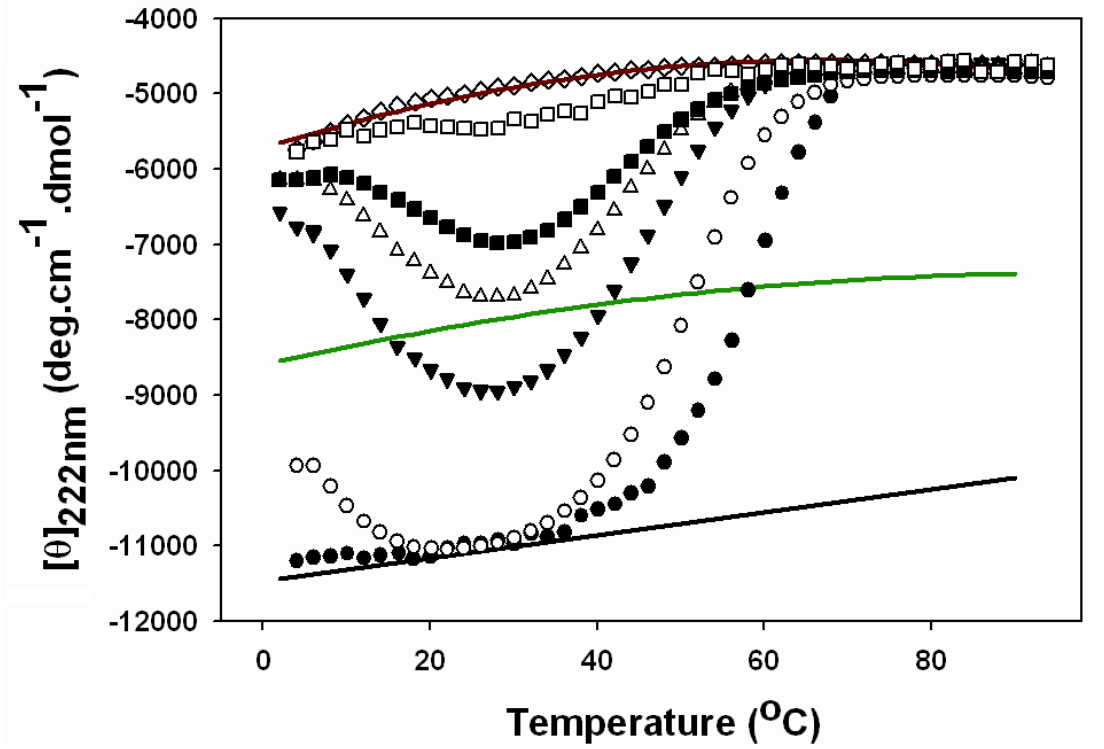


Figure 4-3. CD detected thermal unfolding curves for I98A CTL9 monitored at different pH values. The red curve represents a quadratic fit to the pH 4.7 data and provides an experimental estimate of the unfolded state signal as a function of temperature since the fraction of folded is 0.0 under these conditions. The pH 8.0 data provides an estimate of the CD signal for the fully fold state as the fraction of folded is 1.0 at 25 ^{o}C, pH 8.0. The solid black line represents an extrapolation of the folded baseline. The green curve represents the signal expected for a fraction of folded of 0.50, i.e., the middle point of the transition. (●) pH 8.0; (○), pH 6.6; (▼) pH 6.0; (Δ) pH 5.7; (■) pH 5.6; (□) pH 5.2; (◊) pH 4.7.

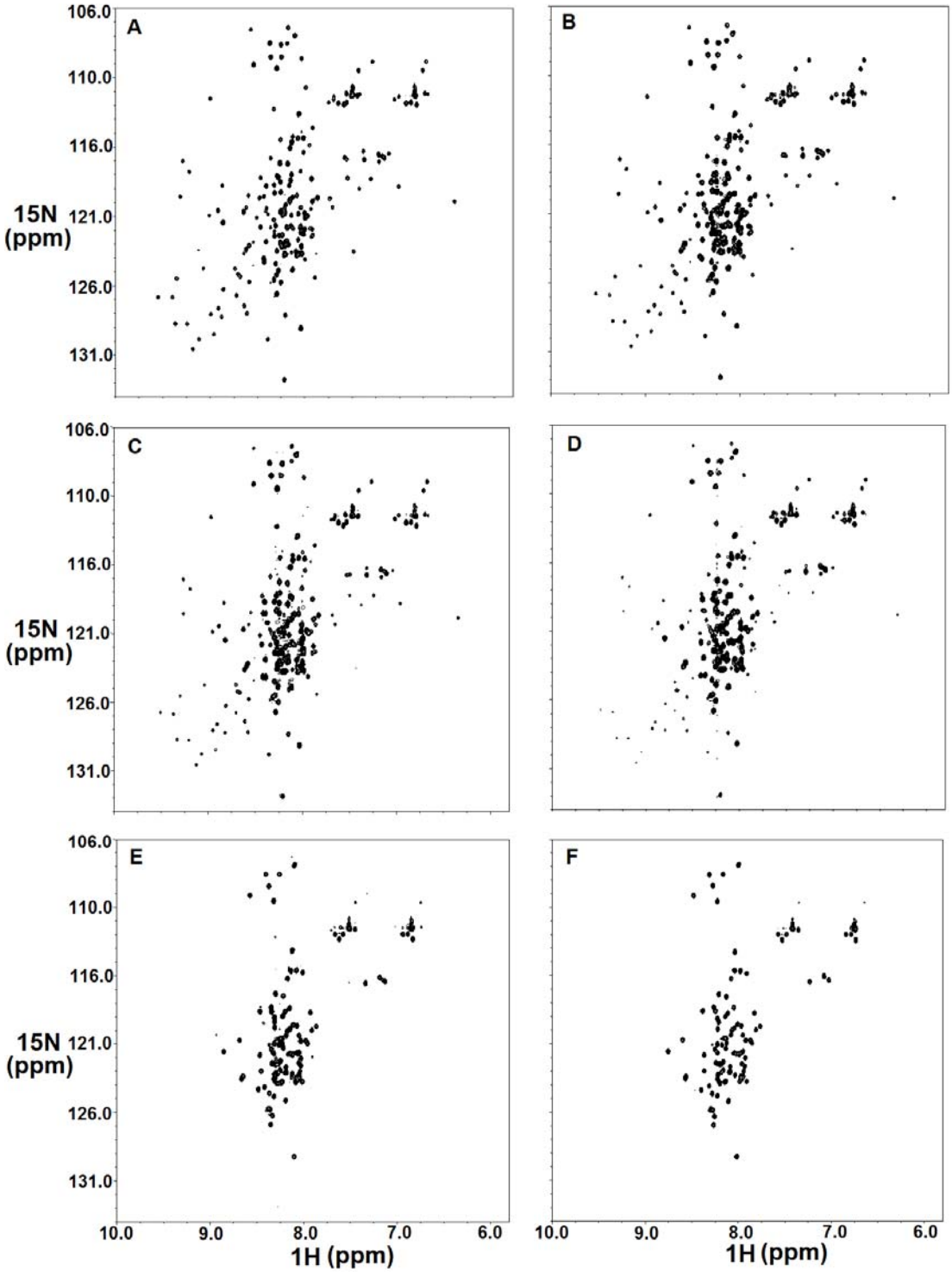


Figure 4-4. ^{1}H-^{15}N HSQC spectra of I98A CTL9 at pH 5.7, in 10 mM NaOAc and 150 mM NaCl. (A) 25 ^{o}C, (B) 21 ^{o}C (C) 18 ^{o}C (D) 15 ^{o}C (E) 11 ^{o}C and (F) 8 ^{o}C.

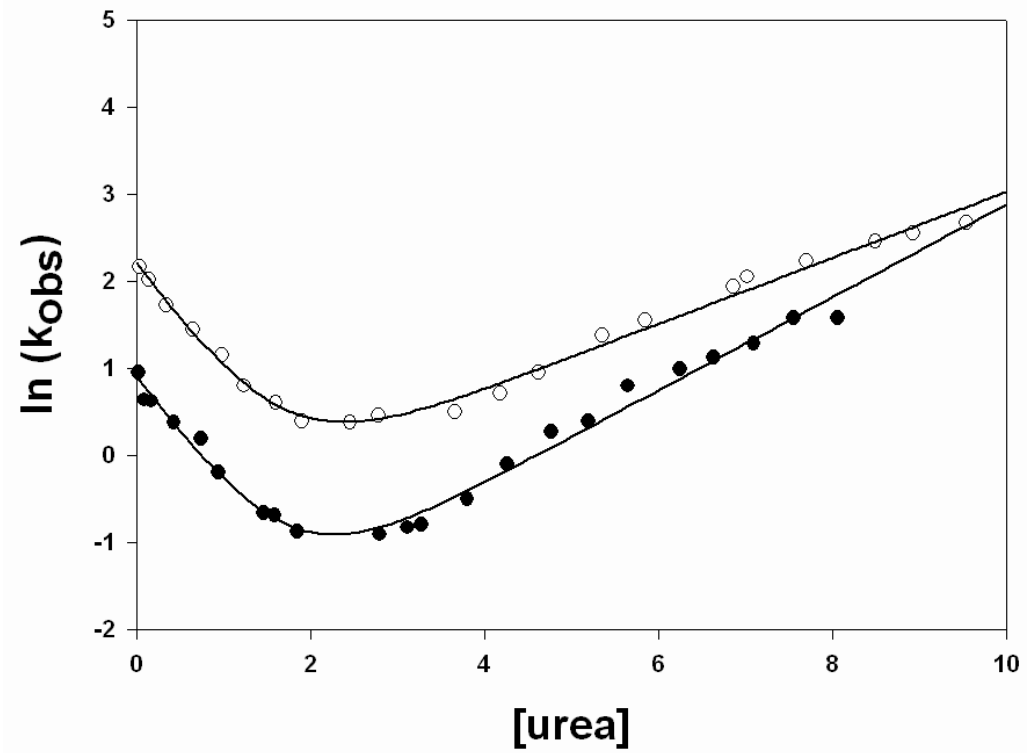


Figure 4-5. Chevron plots of the I98A mutant of CTL9 at pD 6.6, 25 ^{o}C (○) and 15 ^{o}C (●). The continuous line represents the best fit to equation (2).

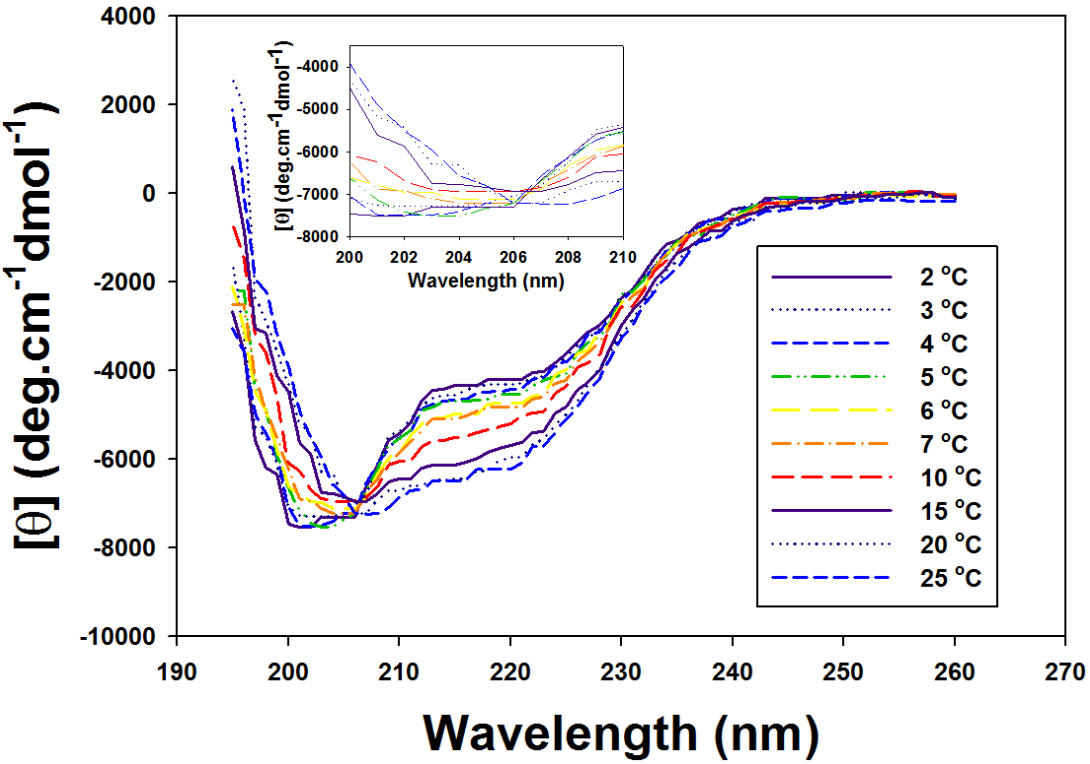


Figure 4-6. Far-UV CD spectra of I98A mutant of CTL9 as a function of temperature, recorded in the temperature range from 2 ^{o}C to 25 ^{o}C at pH 5.7, in 10 mM NaOAc and 150 mM NaCl in H_2O. The inset shows the expansion of the region from 200 nm to 210 nm.

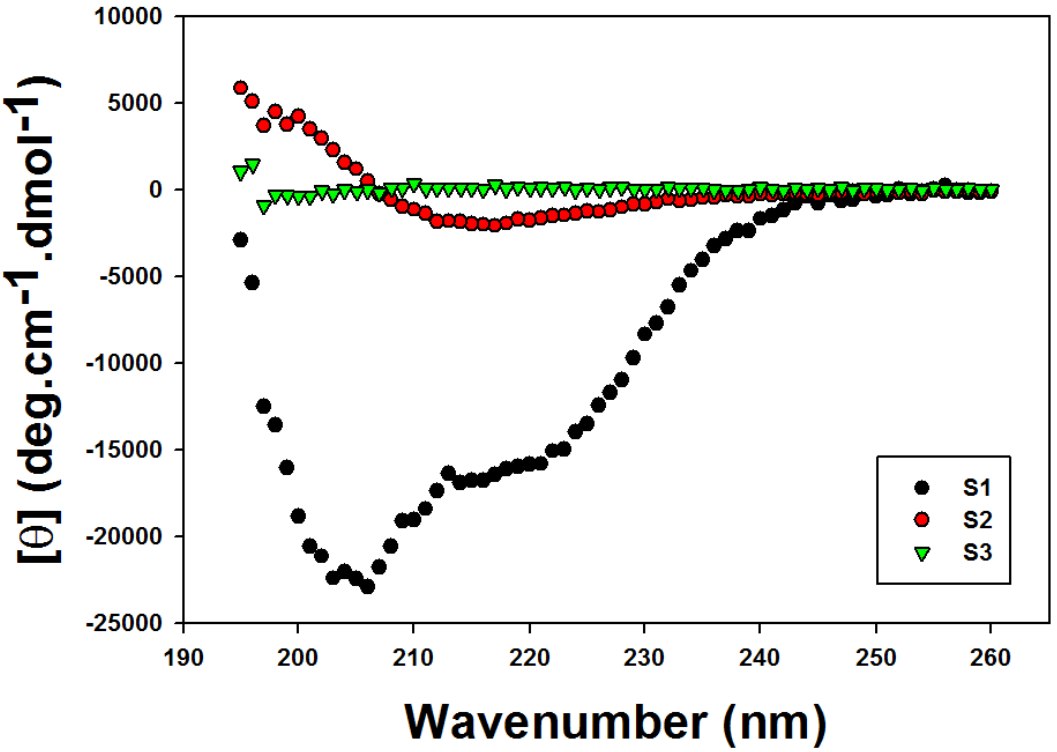


Figure 4-7. SVD of the combined temperature dependent CD spectra of I98A CTL9 at pH 5.7, in 10 mM NaOAc and 150 mM NaCl in H_2O. The three singular vectors with the largest singular values are displayed. The relative amplitudes are 1.00 for S1, 0.14 for S2 and 0.02 for S3.

5. The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure

Cold denaturation is a general property of globular proteins, and the process provides insight into the origins of the cooperativity of protein folding and the nature of partially folded states. Unfortunately, studies of protein cold denaturation have been hindered by the fact that the cold denatured state is normally difficult to experimentally access. Special conditions such as addition of high concentrations of denaturant, encapsulation into reverse micelles, the formation of emulsified solutions, high pressure or extremes of pHs, have been applied, but these can perturb the unfolded state of proteins. The cold denatured state of the C-terminal domain of the ribosomal protein L9 can be populated under native like conditions by taking advantage of a destabilizing point mutation which leads to cold denaturation at temperatures above zero ^{o}C. This state is in slow exchange with the native state on the NMR time scale. Virtually complete backbone ^{15}N, ^{13}C and ^{1}H as well as side-chain $^{13}C_{\beta}$ and $^{1}H_{\beta}$ chemical shift assignments were obtained for the cold denatured state at pH 5.7, 12 ^{o}C. Chemical shift analysis, backbone N-H residual dipolar couplings, amide proton NOEs and R_2 relaxation rates all indicate that the cold denatured state of CTL9 contains significant native like secondary structure, but also contains non-native structure. The regions corresponding to the two native α-helices show a strong tendency to populate helical Φ and Ψ angles. The segment which connects α-helix 2 and β-strand 2 (residues 107-124) in the native state exhibits a significant preference to form non-native helical structure in the cold denatured state. The structure observed in the cold denatured state of the I98A mutant is similar to that observed in the pH 3.8 unfolded state of wild type CTL9 at 25 ^{o}C, suggesting that it is a robust feature of the denatured state ensemble of this protein. The implications for protein folding and for studies of cold denatured states are discussed.

Abbreviations used:
CTL9, the C-terminal domain of the ribosomal protein L9; I98A CTL9, the isoleucine to alanine point mutant of CLT9 at residue 98; NOE, nuclear overhauser effect; PFG NMR, pulsed-field gradient nuclear magnetic resonance; RDC, residual dipolar coupling; SSP, secondary structure propensity.

5.1. Introduction

Cold denaturation is a general property of globular proteins, as predicted by the Gibbs-Helmholtz equation and is believed to be driven by the increased hydration of non-polar groups of proteins at low temperatures as well as by the decrease in hydrophobic interactions within the protein (*95, 98, 100, 187*). Thus the cold induced unfolding of a protein is sometimes thought to proceed via a different mechanism from thermal unfolding. The process of cold denaturation has been studied for a wide range of proteins but structural characterization of the cold unfolded state is less extensive than studies of heat or denaturant induced unfolded states. Investigating the structure of the cold denatured state and the cooperativity of cold denaturation is crucial for a complete understanding of protein folding and cooperativity (*98, 173, 188*) . An increased knowledge of the cold denatured state is also important for understanding the molecular basis for cold adaptation in psychrophilic organisms (*189-190*). Most studies of cold unfolding have focused on thermodynamic aspects of the process or have used relatively low resolution spectroscopic methods.

Unfortunately, the study of cold denaturation is difficult because the midpoint of the cold unfolding transition is typically well below 0 ^{o}C (*98, 100*) . As a consequence studies have been limited to a relatively small number of proteins compared to studies of thermal unfolding. Furthermore the majority of cases require special conditions such as encapsulation (*36, 178, 191*), high pressure (*175-177*), emulsified samples (*180*), extremes of pH (*98, 100, 174, 187*) or the addition of high concentrations of denaturant (*55, 98, 187*) . All of these methods have the potential to offer insight into the structural and thermodynamic aspects of cold denaturation, although some approaches may perturb the system in complicating ways. For example, encapsulation in reverse micelles has been used to study cold unfolding, but our previous NMR diffusion measurements have shown that cold denatured proteins expand as the temperature is lowered (*108*) and other groups have shown that the reverse micelle systems used shed water and contract as the temperature is decreased (*178, 180, 186, 192*). The combination could lead to significant protein micelle interactions at low temperatures. There have been very few reported high resolution studies of the cold denatured state in the absence of encapsulation, denaturant or extremes of pH or pressure (*193*). Here, we report the detailed characterization of the structural and dynamic properties of the cold denatured state using NMR spectroscopy. We use a destabilizing point mutant of the C-terminal domain of the ribosomal protein L9 (CTL9) to populate the cold unfolded state in homogenous solution in the absence of denaturants at near neutral pH. The implications for the folding of CTL9 are discussed as are the implications for the nature of the cold denatured state.

The C-terminal domain of the ribosomal protein L9 (CTL9) is a 92 residue globular protein which adopts a relatively rare fold made up of two α-helices and a three stranded mixed parallel, anti-parallel β-sheet (Figure 5-1A) (*103*). An I98A point-mutation destabilizes the protein and leads to observable cold-denaturation at temperatures above zero ^{o}C in the absence of denaturant or extreme pH values (*108*) (Figure 5-1B and 5-1C). We have characterized the hydrodynamic properties of the cold induced unfolded state of the I98A mutant of CTL9 under native and near native conditions using pulsed field gradient NMR diffusion experiments, and we have shown that the cold denatured state is compact relative to the acid or urea unfolded states but expands as the temperature is lowered (*108*).

5.2. Materials and Methods

5.2.1. Protein expression and purification

^{15}N-labeled CTL9 and ^{13}C, ^{15}N-labeled CTL9 and the labeled I98A mutant were overexpressed and purified as previously described (*32, 78*). The identity of the proteins was confirmed using MALDI-TOF mass spectroscopy and the purity was checked using analytical HPLC. The yield of pure protein was 30 mg per liter of M9 minimal media.

5.2.2. Circular dichroism spectroscopy

CD experiments were carried out using a Chirascan CD spectrometer. The protein was dissolved in 20 mM acetate, 100 mM NaCl buffer, pH 5.7 at a concentration of about 18 μM. Thermal denaturation experiments were performed by monitoring the ellipticity at 222 nm with a heating rate of 1 ^{o}C/min. Data were collected from 2 ^{o}C to 94 ^{o}C, with a 2 ^{o}C increment. The experiments were repeated 3 times with independently prepared samples to access the reproducibility of the results and the experimental uncertainties. Additional CD melting experiments were conducted as a function of pH between pH 4.8 and pH 8.0.

5.2.3. Stopped-flow fluorescence

Direct measurement of the exchange rate between the cold denatured state and the folded state at pH 5.7, 12^{o}C was performed by a fluorescence detected pH jump experiments as described in chapter 4.

5.2.4. Nuclear magnetic resonance experiments

Protein samples for NMR experiments were prepared in 90% H_2O and 10% D_2O at pH 5.7, with 20 mM sodium acetate and 100 mM NaCl, unless otherwise specified. The protein concentration was about 1 mM. All NMR experiments for the cold denatured state of I98A CTL9 were recorded on 800 or 900 MHz Bruker spectrometers equipped with cryogenic probes at the New York Structural Biology Center. Triple resonance experiments for establishment of the assignments for wild-type CTL9 were performed on a 600 MHz Varian spectrometer at the University of Connecticut Health Center. The ^{1}H dimension was centered at the water resonance and the ^{15}N offset frequency was set to 118.0 ppm for all heteronuclear NMR experiments.

^{15}N-^{1}H correlated heteronuclear single coherence (HSQC) spectra for the I98A mutant were recorded using 1024 x 512 complex points with 8 scans per increment and spectral widths of 9920.6 Hz and 2189.6 Hz for the ^{1}H and ^{15}N dimensions, respectively. The HSQC spectra of wild-type CTL9 sample had spectral widths of 6600.7 Hz (^{1}H) and 1920.0 Hz (^{15}N), respectively.

The HNCO/HNCACO (*153-155*) and the HNCACB (*156-157*) experiments were performed to enable sequence-specific backbone and sidechain $^{13}C_{\beta}$ assignments. The HBHACONH (*157-158*) experiment was carried out to obtain sidechain $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ assignments. The watergate sequence was employed to suppress water signals in all triple resonance experiments. All spectra were processed using NMRPipe software (*135*), and chemical shift assignments were made using NMRViewJ (*136*). All the chemical shifts were referenced to the absolute frequency of DSS at 0.00 ppm. The random coil chemical shift values of Wishart were used to calculate the secondary chemical shifts (*87*). SSP scores were calculated based on $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and $^{1}H_{\alpha}$ shifts, using the software provided by Professor Forman-Kay at her website (http://pound.med.utoronto.ca/software.html) (*139*).

A 3D ^{15}N HSQC-NOESY-HSQC experiment was performed on a ^{15}N-labeled sample of I98A CTL9. The spectrum was acquired with spectral widths of 8012.8 (NH),

2108.8 (^{15}N) and 2108.8 (^{15}N) Hz and in a data matrix of 1024 x 128 x 138 complex points, with a mixing time of 300 ms. Coherence selection with gradients and sensitivity enhancement were used, and water suppression was achieved using a water flip back pulse sequence. NOESY peak volumes were measured using NMRViewJ. The NOE peak volumes were normalized as the ratio of the *NN*(i,i+1) cross peak to the *NN*(i, i) diagonal peaks to compensate for any potential relaxation effects on the NOE intensities.

^{15}N-^{1}H one bond residual dipolar couplings (RDC) were measured in weakly aligned bicelle media. The bicelles were prepared by mixing 450 μL of NMR buffer with 50 μL of pentaethylene glycol octylether (C8E5). The mixture was gently vortexed until the solution became clear. Then 1-octanol was added in 2 μL increments for a total additional volume of 10 to 11 μL. The bicelles were left at 10 ^{o}C overnight. Then the protein solution was added at a ratio of 1:1 to the bicelle solution and gently vortexed. The alignment of the protein with the bicelles was checked by measuring the ^{2}H quadrupolar splitting of the water line. A ^{15}N-^{1}H HSQC experiment was performed on the protein sample in the bicelles. There was no significant change in either the chemical shift or peak intensity observed in any of the resonances, indicating the bicelles do not perturb the protein conformation significantly. To measure the ^{15}N-^{1}H RDCs, 2D inphase and antiphase-HSQC (IPAP-HSQC) spectra (*194*) were recorded for the sample in the bicelles and in isotropic buffer solutions. The IPAP-HSQC spectra were acquired with 2048 X 512 complex points and with spectral widths of 8012.8 Hz and 2270.7 Hz for the ^{1}H and ^{15}N dimension, respectively. The spectra were processed using NMRPipe software and the ^{15}N-^{1}H splittings were measured using NMRViewJ. The RDCs were calculated as the difference between the scalar coupling splitting in isotropic solution and the observed splitting in the anisotropic solution.

Transverse relaxation (R_2) experiments were performed on a ^{15}N-labeled sample as described using a Carr Purcell Meiboom Gill (CPMG) sequence (*89*). The pulse spacing in the CPMG sequence was 0.9 milliseconds. 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 a recycle delay of 3 s. To estimate the uncertainty, duplicate spectra were acquired at 50.88, 101.76 and 152.64 ms delay times. Each spectrum was collected with 4 scans and 1024 x 256 complex points. The spectral widths were 8012.8 Hz (^{1}H) and 2027.7 (^{15}N). R_2 rates were determined by fitting the NMR resonance intensities as a function of delay times to a single exponential decay using NMRViewJ. The expected intrinsic R_2 rates for a completely unstructured protein were obtained by fitting the experimental R_2 rates to the phenomenological model proposed by Schwalbe and coworkers (*90*).

5.3. Results

5.3.1. The I98A mutant of CTL9 adopts the same structure as wild-type CTL9 but undergoes cold denaturation above ^{o}C

Our previous thermodynamic studies indicated that replacing the hydrophobic core residue I98 in CTL9 by an alanine causes a decrease in $\Delta H^{0}(T_m)$ and an increase in $\Delta C_p{}^{0}$, thus the temperature of the mid point of cold denaturation increases accordingly (*108*). At pH values near 6.0 cold denaturation was observed at temperatures above 0 ^{o}C by both CD monitored temperature melting experiments and by NMR monitored melting studies (Figure 5-1B and 5-1C). As shown in Figure 5-1C, resonances from both the native and denatured state can be detected at 25 ^{o}C. As the temperature decreases, the resonances of

the cold denatured state increases while the peaks from the native state diminish gradually. Similar to the wild-type protein, the stability of I98A CTL9 is strongly dependent on pH due to the protonation of one or more of the three histidine side-chains (*32*). At pH 5.7, the cold denatured state is 80% populated at 12 ^{o}C (Figure 4-3), making NMR studies of the cold denatured state feasible. Thus, we chose pH 5.7, 12 ^{o}C as the experimental condition to explore the structural and dynamic properties of the cold denatured state of I98A CTL9.

The HSQC spectrum of the cold denatured I98A CTL9 is typical for an unfolded protein (Figure 5-2), and one major set of peaks are observed, indicating that the cold denatured state is the dominant conformer, a result which is in agreement with the 1D NMR experiments. Peaks from the native state are observed at lower contour levels. All of the peaks are sharp as expected for a monomeric protein, which is not experiencing exchange broadening. We used pH jump fluorescence detected stopped flow measurements to directly measure the exchange rate at 12 ^{o}C. The exchange rate is 0.5 sec^{-1} (Figure 5-3), confirming that the system is in slow exchange on the chemical shift time scale.

Using triple resonance HNCACB, HNCO and HNCACO experiments, virtually complete backbone ^{13}C, ^{1}H and ^{15}N assignments were obtained. Sidechain $^{1}H_{\alpha}$ and $^{1}H_{\beta}$ were also made using the HBHACONH experiment. The assignments have been deposited in the Biological Magnetic Resonance data Bank (BMRB).

5.3.2. Chemical shift analysis provides evidence of native and non-native secondary structure in the cold denatured state

Chemical shifts are valuable probes of protein structure because of the strong correlation between secondary chemical shifts, i.e., the deviations of observed chemical shifts from random coil chemical shifts, with protein structure. $^{13}C_{\alpha}$ and ^{13}CO shifts are shifted downfield in α-helices and upfield in β-sheets, while the opposite trend is usually observed for $^{1}H_{\alpha}$ and $^{13}C_{\beta}$ shifts (*83, 86, 161*). The $^{13}C_{\alpha}$, ^{13}CO, $^{1}H_{\alpha}$ and $^{13}C_{\beta}$ secondary shifts of the cold denatured state of I98A CTL9 were calculated using the random coil values reported by Wishart (*87*). The data are presented in Table 5-1 and Figure 5-4. The various observed trends are broadly self-consistent, and indicate that the two α-helical regions (residues 58-74 and residues 95-106) in the native state have a strong tendency to form α-helical structure in the cold denatured state. In addition, the loop region which connects α-helix 2 and β-strand 2, i.e., residues 107-124 in the native state, also exhibits a clear trend to populate helical Φ and Ψ angles. This is of interest because it implies that non-native structure is formed.

Individual secondary shifts are sensitive to protein secondary structure but their predictive value strongly depends on the choice of chemical shift standard. Combinations of secondary shift data can often provide a more accurate picture of the structural propensities for unfolded and partially unfolded proteins. We calculated the difference between $^{13}C_{\alpha}$ secondary shifts and $^{13}C_{\beta}$ secondary shifts, as plotted in Figure 5-5. The figure compares the $\Delta\delta C_{\alpha}$-$\Delta\delta C_{\beta}$ values for the cold denatured state of the I98A mutant (Figure 5-5A) and the native state of the wild type CTL9 (Figure 5-5B). Positive $\Delta\delta C_{\alpha}$-$\Delta\delta C_{\beta}$ values indicate α-helical structure and negative values indicate a propensity to adopt β-strand structure (*139*). Positive $\Delta\delta C_{\alpha}$-$\Delta\delta C_{\beta}$ values are detected for the two helical regions as well as for the loop between α-helix 2 and β-strand 2, with average values of

1.26, 1.01 and 0.91, respectively. For comparison, the values for folded wild-type CTL9 under the same conditions are 3.47, 2.21 and 0.21 (Figure 5-5B and Table 5-2), respectively. Negative $\Delta\delta C_\alpha$-$\Delta\delta C_\beta$ values, with an average of -0.17, are observed in residues 78-82, which reside in the first β-strand in the native state, while the average is -1.79 in the wild-type CTL9 under the same conditions.

Forman-Kay and coworkers have developed the SSP (secondary structure propensity) algorithm to detect secondary structure propensities in partially unfolded and natively disordered proteins (*139*). The SSP score is the weighted average of the chemical shifts from different nuclei in a given residue, with the relative weighting reflecting the sensitivity of different secondary shifts to structure. A SSP score of 1.0 suggests a fully formed α-helix, while a value of -1.0 suggests a fully formed β-strand. Figure 5-6 displays the SSP scores of the cold denatured state of I98A at pH 5.7, 12 ^{o}C and those of the wild-type CTL9 under the same conditions. Wild-type CTL9 is well folded under these conditions. The cold denatured state of I98A, in comparison, is less structured and as expected, exhibits SSP scores with much smaller amplitude. Nevertheless, some obvious structural trends are detected. Positive values are observed for α-helix 1 and α-helix 2. The average is 0.33±0.08 for α-helix 1 and 0.25±0.06 for α-helix 2. Large positive SSP scores are also found in the region which connects α-helix 2 and β-strand 2 (residues 107-124), with an average of 0.20±0.11. For comparison, the values for folded state of the wild-type CTL9 are 0.91±0.28, 0.78±0.31 and 0.09±0.50 respectively. Negative SSP scores are observed in the cold denatured state for residues 78-83 and residues 128-133, which are located in the first and second β-strand respectively, in the native state,. The average SSP values for these two segments are -0.17±0.10 and -0.18±0.06. The corresponding values for folded wild-type CTL9 are -0.51±0.32 and -0.42±0.20 under the same conditions. The SSP analysis is in excellent agreement with the secondary shifts analysis, which argues that the cold denatured state of I98A CTL9 contains both native like and non-native α-helical as well as a tendency to sample native like β-strand or extended structures.

5.3.3. Residual dipolar couplings confirm α-helical secondary structure in the cold denatured state

Residual dipolar couplings (RDC) are playing an increasingly important role in structural determination of macromolecules (*195-196*). For well folded globular proteins, the magnitude and sign of the RDCs depend on the orientation of the bond vectors relative to the alignment tensor of the entire molecule. In contrast, unfolded proteins generally can not be described by a global alignment tensor. Instead, the polypeptide chain can be viewed as behaving like a collection of short segments of polypeptide in which each short segment transiently aligns relative to the external magnetic field. The length of the segment, often termed the persistence length, reflects the number of amino acid residues within which the polypeptide chain has a correlated orientation. Previous work has suggested that the NH bond vectors adopt alignments that are approximately perpendicular for α-helical and β-strand or extended structure. Thus the NH RDCs in independently aligning α-helices and β-strand have opposite signs since the NH vector is aligned parallel to the long axis of an α-helix but perpendicular to the long axis of a β-strand (*92-93, 197-198*). As a result, NH RDCs can report on residual secondary structure in unfolded proteins. In contrast, the RDCs of highly unfolded proteins usually have the

same sign. Recent work has shown that RDCs for denatured states with little or no residual structure can be reasonably well predicted using ensembles constructed from dihedral angle choices derived from certain regions of the protein data bank(*199-200*). The predicted RDCs are dominated by those that come from highly extended conformations, since these align most strongly with the field and thus explain the observation that highly unfolded proteins exhibit strong RDCs with mostly the same sign. If there is local residual structure such as helices in an otherwise extended polypeptide, then the alignment of the NH bond vectors in the helical segments will necessarily be perpendicular to that of the NH bond vectors in the extended regions. Irrespective of whether one wishes to describe this as local alignment of persistence-length segments or alignment of highly extended conformations, the end result is that RDCs are well suited for probing residual helical structure in unfolded proteins. In the current study, NH RDCs of the cold denatured state of I98A CTL9 were measured in weakly aligned bicelles. The HSQC spectrum obtained in the bicelle media (data not shown) are identical to the one collected in isotropic solution, indicating that the alignment media does not significantly perturb the conformation. Figure 5-7A displays the NH RDCs measured for the wild-type CTL9 at pH 5.7, 12 ^{o}C. Although the globular wild-type protein aligns in the bicelle medium as a rigid structure, one can note that the NH RDCs are positive for α-helix 1 and negative in α-helix 2, consistent with the native structure in which the two helices are oriented perpendicular to each other. All of the NH RDCs are of the same sign (negative) in the β-strand regions, indicating that the β-strands are oriented either parallel or anti-parallel to each other. The RDC data of the native state of CTL9 agrees well with the crystal structure. The NH-RDCs for the cold denatured state of I98A CTL9 are plotted in Figure 5-7B. The segments which are predicted to populate α-helical structure based on the chemical shift analysis, including the loop connecting α-helix 2 and β-strand 2, exhibit RDCs which have opposite signs relative to the β-strand regions or extended regions. The three β-strand segments display negative RDCs. It is worth noting that β-strand 3 (residues 138-147) displays significant negative RDCs in the cold denatured state in spite of the small SSP scores in this region. A plausible explanation is that the NH vectors are aligned parallel in β-strands and extended structure in unfolded proteins (*93*). Other than the β-strand 3 region, the pattern of the RDC data is in excellent agreement with the SSP scores and the RDC data offers additional evidence that there is significant helical structure in the cold denatured state. It is noteworthy that all RDCs have much smaller amplitude in the cold denatured state than in the native state. Since the experimentally measured RDCs are population weighted averages of an ensemble of conformations, the small magnitudes of the RDC reveal that the residual secondary structure in the cold denatured state is only partially developed. Note that the measured RDCs for the cold denatured state do not include any contributions from averaging from the native structure since the native and the cold denatured states are in slow exchange.

5.3.4. Amide proton NOEs are consistent with native like structure in the cold denatured state

Proton NOEs provide valuable structural probes: Short-range and medium range NOEs are widely used to delineate protein secondary structure and long-range NOEs are particularly useful for defining the global topology of proteins (*201*). However, measurements of NOEs in unfolded proteins are quite difficult due to the flexible nature

of the polypeptide chain in the unfolded state ensemble and the high degeneracy of the resonances. Nevertheless, sufficient spectral resolution can be obtained using the ^{15}N-HSQC-NOESY-HSQC experiment, provided that the population of a conformation with a defined secondary structure is high enough. Sequential amide proton NOEs, which are often used to distinguish α-helices from β-strand and extended structure, are detected throughout the protein in the cold denatured state (Figure 5-8). The two α-helical regions as well as the loop region connecting α-helix 2 and β-strand 2, display sequential amide NOEs which have higher intensities than the NOEs observed for other segments of the polypeptide chain. The average intensities of the normalized NOEs are 1.7 times as large in these three regions than are observed for other residues. The observed NOEs are consistent with the chemical shift and RDC measurements however there is the possibility that some of the observed effects could be due to transferred NOEs from the native state. Transferred NOE experiments are typically performed for systems in fast exchange (*202-203*), however they have been observed for proteins in slow exchange (*204-206*). In the present case the NOE mixing time was 300 milliseconds and the measured exchange rate between the cold denatured state and the native state is 0.5 sec^{-1}, corresponds to time constant of 2 seconds thus transferred NOE effects may contribute. The key point, however, is that the data is consistent with the chemical shift and RDC data and those measurements will not be affected by the slow exchange.

5.3.5. ^{15}N transverse relaxation rate analysis of the cold denatured state

Relaxation rate measurements have been widely used to study the structure and dynamics of unfolded proteins (*207-208*). Transverse ^{15}N relaxation rates (R_2) are particularly useful because they are very sensitive probes of deviations from random coil behavior in unfolded proteins. For a completely unstructured peptide with only local interactions, the ^{15}N R_2 relaxation rates are expected to be uniform in the center of the polypeptide chain and slightly smaller at the ends (*90-91*). The observation of significantly larger R_2 rates is thought to reflect the formation of local residual structure such as hydrophobic clusters (*63, 83, 90-91*). Figure 5-9 plots the measured relaxation rates and the relaxation rates calculated based on the Schwalbe phenomenological model (*37)*. The slow exchange between the folded and the cold denatured state will not affect the measured R_2 rates. The calculated relaxation rates follow the expected inverted U-shaped profile. The fit yields an apparent persistence length of 8 residues, which is consistent with previous estimates (*182*) and supports the analysis of the RDC data. Not surprisingly, the native state displays large systematic deviations from the values calculated using the Schwalbe random coil model. The ^{15}N relaxation rates for the cold denatured state exhibit smaller but still significant positive deviations from the random coil model. In particular, the largest deviations are found in the two α-helical regions as well as in the loop connecting α-helix 2 and β-strand. No significant deviations from the model random coil relaxation rates are found in the β-strand regions, suggesting that the α-helical structure is better formed in the cold denatured state than the β-strand structure.

5.4. Discussion

Studies of the cold denatured state have often relied on special conditions such as addition of denaturants, high pressure, encapsulation, emulsified samples or extremes of pH, under which the structural and dynamic properties of the protein may be perturbed.

Nevertheless these investigations have provided valuable information about unfolded proteins. There have been a few reported studies of cold denaturation under near native conditions (*173, 193, 209-210*), however, to the best of our knowledge, the current study provides the first detailed residue specific characterization of the cold denatured state of a protein under near physiological conditions. We have obtained essentially complete ^{1}H, ^{13}C and ^{15}N backbone assignments as well as $^{13}C_{\beta}$ and $^{1}H_{\beta}$ sidechain assignments for the cold denatured state of I98A CTL9 at pH 5.7 and 12 ^{o}C. The chemical shift analysis, NOE analysis, RDC data and transverse relaxation rates all yield a consistent picture. Residues in the two α-helical regions in the native state have a strong tendency to sample the α-helical region of the Ramachandran plot. Residues 107-124, which form a short 3_{10} helix and the connecting loop linking α-helix 2 and β-strand 2 in the native state, also preferentially form α-helical structure in the cold denatured state. The R_2 values of these three segments exhibit deviations from the phenomenological model for R_2 values in fully unfolded states. Previous studies of the structure of cold denatured states, particularly pressure dependent studies, have suggested that cold denatured states either contain significant residual structure, that often mimics early intermediates which are populated during the kinetic refolding pathway, or are largely unstructured (*177, 211-213*). The cold denatured state of CTL9 exhibit intermediate behavior between these two extremes. The folding of CTL9 has been well documented to be two-state. The work presented here provides additional evidence that the protein folds from a state which contains significant residual secondary structure,

The observed secondary structure and the deviations of the R_2 values from the random coil model are consistent with previous hydrodynamic measurements, which have demonstrated that the cold denatured state of the I98A mutant is more compact than the highly denatured pH 2.0 state or the urea denatured state of the wild-type (*108*). NMR characterization of the acid denatured state and the urea denatured state revealed that significantly less residual structure is present.

The cold denatured state of the I98A mutant is more similar to the pH 3.8 unfolded state of the wild type CTL9, both are compact and both contain native and non-native secondary structure in the same regions of the chain. Indeed, there is a strong correlation between the SSP scores of these two unfolded states, with a r^2 value of 0.84 (p-value$<10^{-4}$) (Figure 5-10). The similarity between these two unfolded states of CTL9 strongly suggest that the residual structure observed is robust and is not an artifact of the conditions used to populate the unfolded states. It is difficult to experimentally detect when a mutation alter the unfolded state structure or energetics, however, the observation of similar helical structural propensities in the cold denatured state of the I98A mutant and the pH 3.8 unfolded state of the wild-type (*182*) argues that the isoleucine to alanine mutation at residue 98 may have little impact on the unfolded state of the protein. Thus the native and non-native structure observed in the I98A mutant is highly likely to be formed in the unfolded state which is populated under refolding conditions.

Nascent residual structure in the unfolded state can reduce the conformational search in the early stages of protein folding. As a consequence, structural information about the unfolded state of CTL9 may have important implications for the folding of this protein (*55*). A previously reported φ-value analysis of the transition state for the folding of CTL9 revealed small φ-values for the α-helices and somewhat larger but still modest φ-values in β-strand 2, β-strand 3 and the loop which connects the two strands (*129*). φ-

value analysis reports on the development of interactions in the transition state relative to the unfolded state and given the significant helical structure in the unfolded state of CTL9, the transition state for the folding of CTL9 likely contains significant helical structure as well. One should also bear in mind that a mutation which destabilizes the unfolded state may decrease the φ-value relative to the value which would be measured for the case where the mutation does not affect unfolded state energetics (*50*).

Table 5-1. Average secondary shifts and SSP values in secondary structural elements of I98A CTL9 at pH 5.7, 12 ºC

	$^{1}H_{\alpha}$	$^{13}C_{\alpha}$	$^{13}C_{\beta}$	^{13}CO	$^{13}C_{\alpha}$-$^{13}C_{\beta}$	SSP
Helix 1	**-0.10**	**1.23**	**-0.03**	**0.76**	**1.26**	**0.33**
Helix 2	**-0.08**	**0.96**	**-0.05**	**0.83**	**1.01**	**0.25**
Loop	**-0.08**	**0.87**	**-0.04**	**0.43**	**0.91**	**0.20**
Strand 1	**-0.01**	**0.06**	**0.23**	**-0.65**	**-0.17**	**-0.17**
Strand 2	**-0.03**	**0.27**	**0.23**	**-0.83**	**0.04**	**-0.18**
Strand 3	**-0.01**	**0.23**	**0.26**	**-0.33**	**-0.03**	**0.00**

Table 5-2. Average secondary shifts and SSP values in secondary structural elements of wild-type CTL9 at pH 5.7, 12 ºC

	$^{1}H_{\alpha}$	$^{13}C_{\alpha}$	$^{13}C_{\beta}$	^{13}CO	$^{13}C_{\alpha}$-$^{13}C_{\beta}$	SSP
Helix 1	**-0.35**	**2.80**	**-0.67**	**1.95**	**3.47**	**0.91**
Helix 2	**-0.28**	**2.91**	**0.70**	**1.90**	**2.21**	**0.78**
Loop	**-0.09**	**0.88**	**0.67**	**0.08**	**0.21**	**0.09**
Strand 1	**0.16**	**-0.51**	**1.28**	**-0.89**	**-1.79**	**-0.51**
Strand 2	**0.50**	**0.04**	**1.57**	**-0.98**	**-1.53**	**-0.42**
Strand 3	**0.48**	**-1.05**	**1.30**	**-1.51**	**-2.35**	**-0.28**

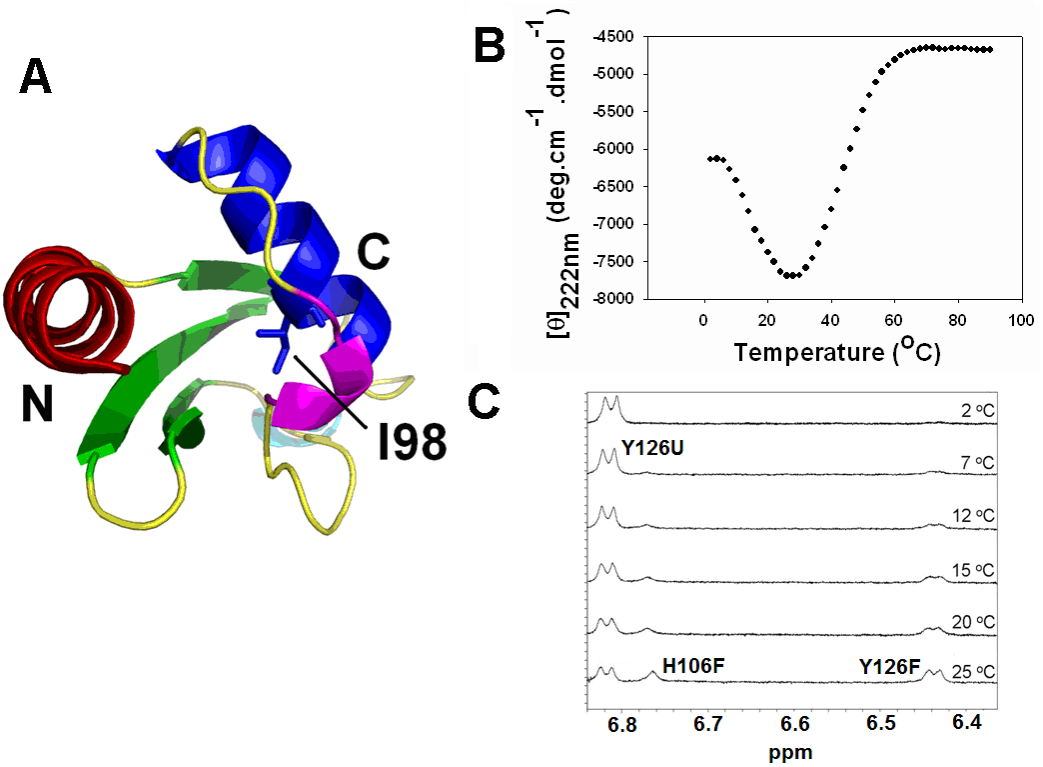


Figure 5-1. The I98A mutant undergoes cold denaturation. (A) Ribbon diagram of CTL9 (residues 58-149 of protein L9), Protein data bank entry 1DIV. The hydrophobic core residue I98 as well as the N and C-termini are labeled. The ribbon diagram was made using PyMol. (B) Thermal denaturation curve of I98A CTL9 monitored by CD at 222 nm in H_2O at pH 5.7. (C) 1D ^{1}H-NMR spectra of I98A CTL9 at different temperatures in 100% D_2O at pD 6.0 (uncorrected pH meter reading). Only the aromatic region is shown, resonance assignments are labeled.

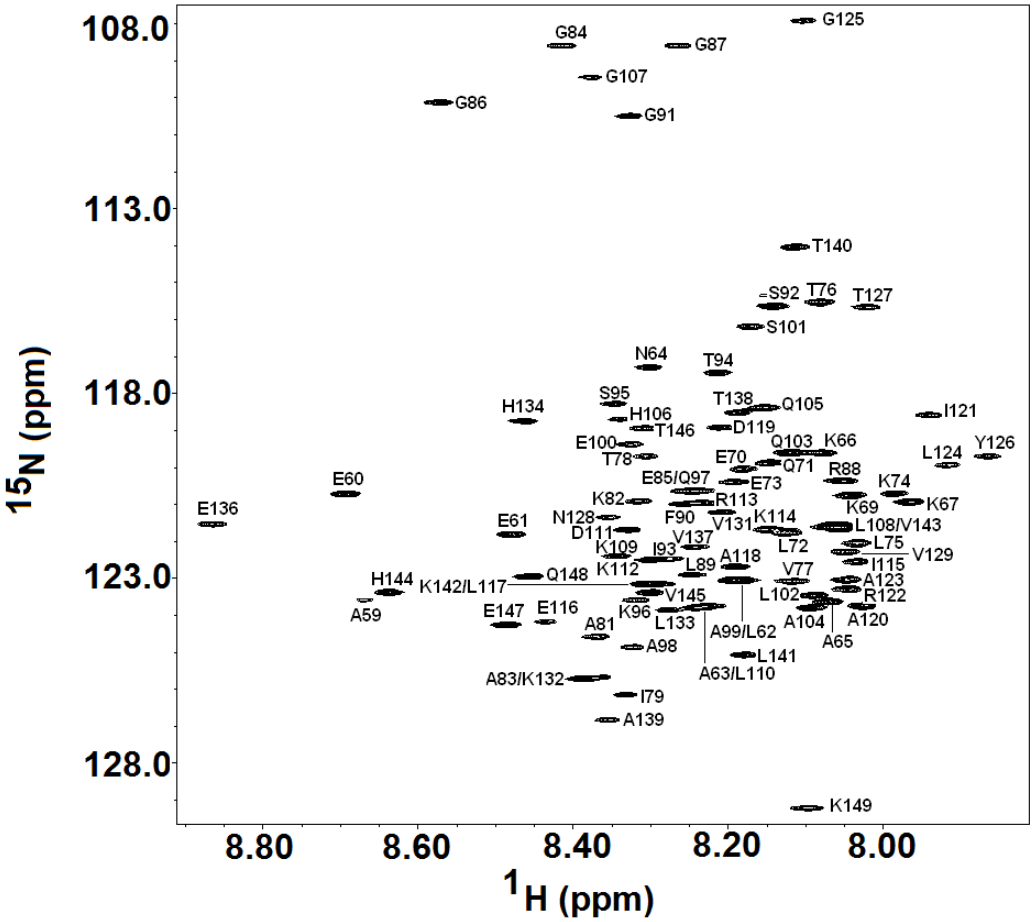


Figure 5-2. ^{15}N-HSQC spectrum of the cold denatured state of I98A CTL9 with the assignments indicated. The spectrum was recorded at pH 5.7, 12 ^{o}C.

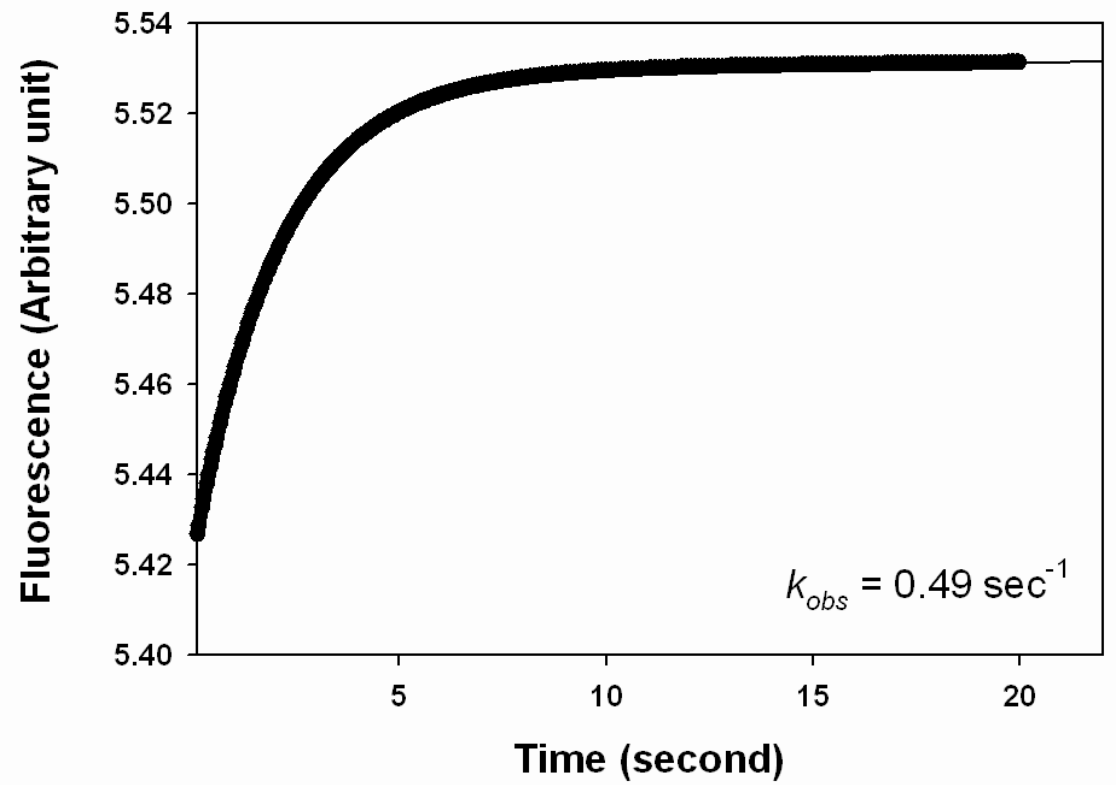


Figure 5-3. The time course of the change in fluorescence which is due to Tyr126. The curve is fit to a double exponential decay. A minor slow phase which corresponds to cis-trans proline isomerization in the unfolded state has a rate constant of 0.037 sec^{-1}. The faster phase corresponds to folding and has a first order rate constant, $k_{obs} = k_{folding} + k_{unfolding}$, of 0.49 sec^{-1}.

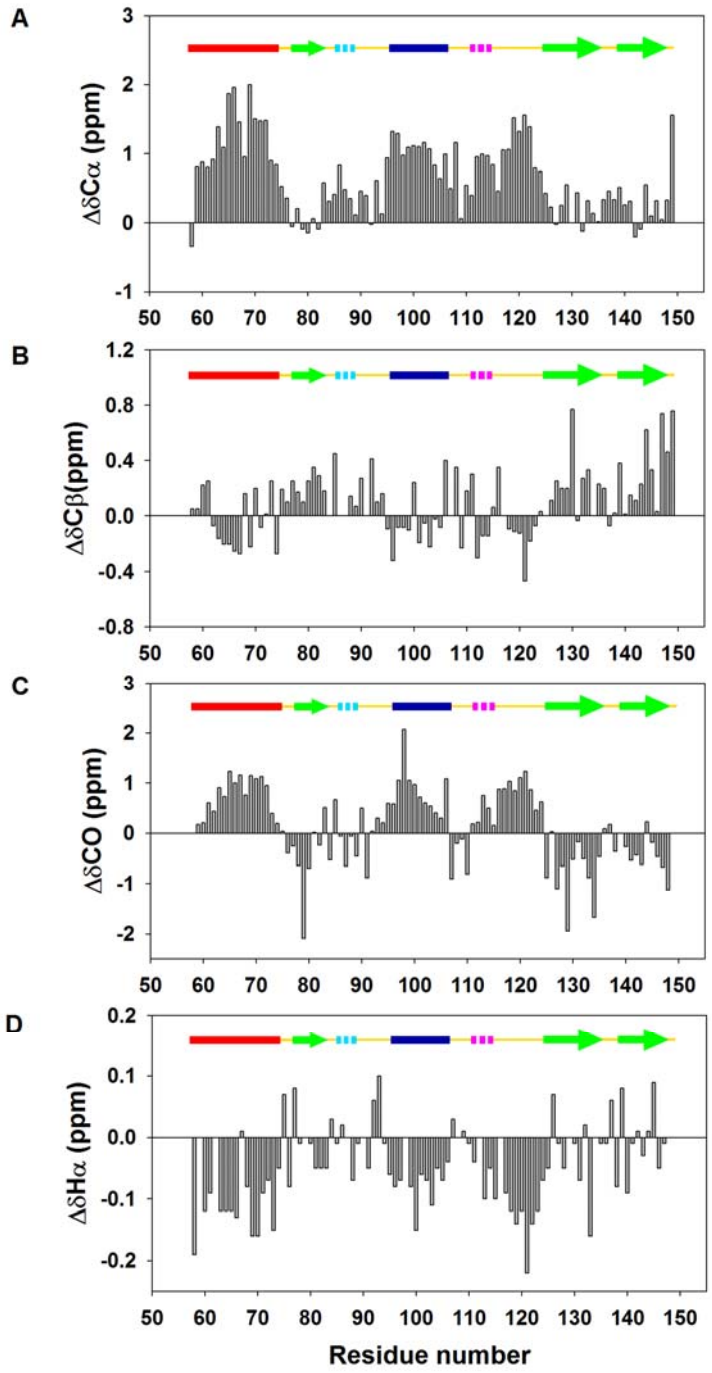


Figure 5-4. Plots of (A) $^{13}C_{\alpha}$ secondary shifts, (B) $^{13}C_{\beta}$ secondary shifts, (C) ^{13}CO secondary shifts and (D) $^{1}H_{\alpha}$ secondary shifts for the cold denatured state of I98A CTL9. Assignments were obtained at pH 5.7, 12 ^{o}C . A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

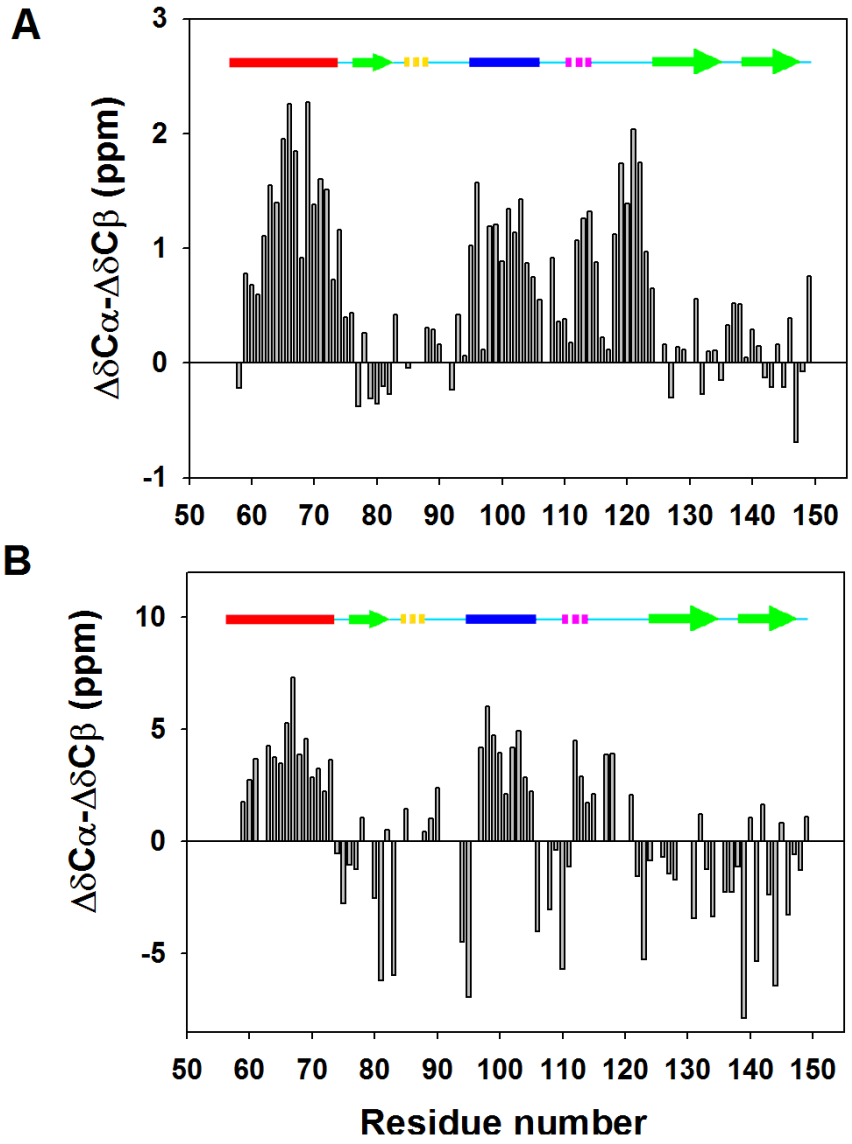


Figure 5-5. Plots of the difference in the $^{13}C_{\alpha}$ secondary shifts and the $^{13}C_{\beta}$ secondary shifts for (A) The native state of wild-type CTL9, (B) the cold denatured state of I98A CTL9. Note that the scale is different in panel A and panel B. Assignments were obtained at pH 5.7, 12 ^{o}C. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel, with solid bars, hashed bars, and open arrows indicating alpha-helix, 3-10 helix and beta-strand structure, respectively. The color coding corresponds to that used in Figure 5-1.

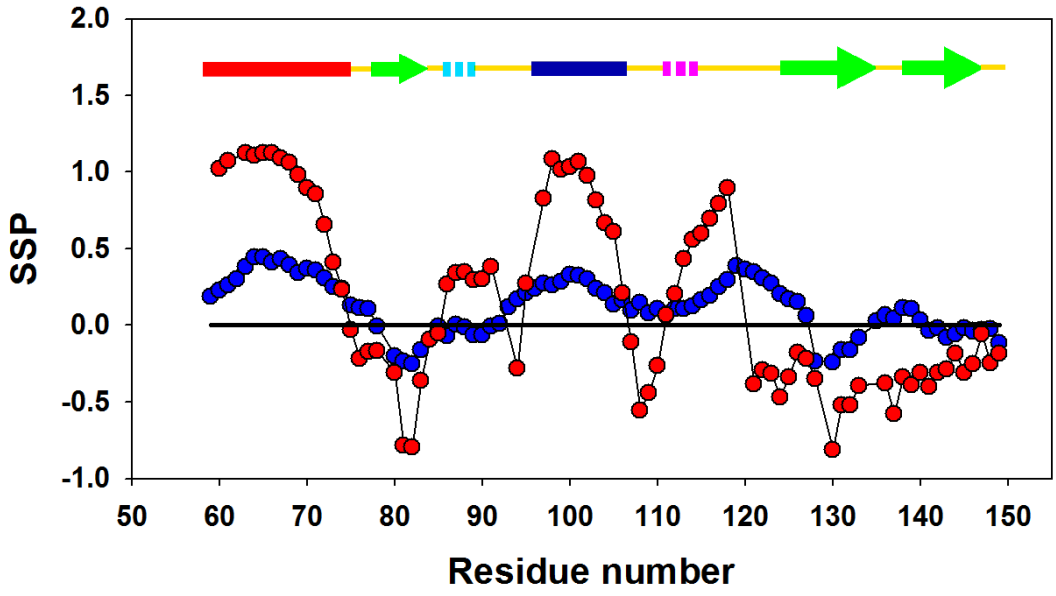


Figure 5-6. SSP analysis of wild-type CTL9 (●) and I98A CTL9 (●) at pH 5.7, 12 ^{o}C . A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

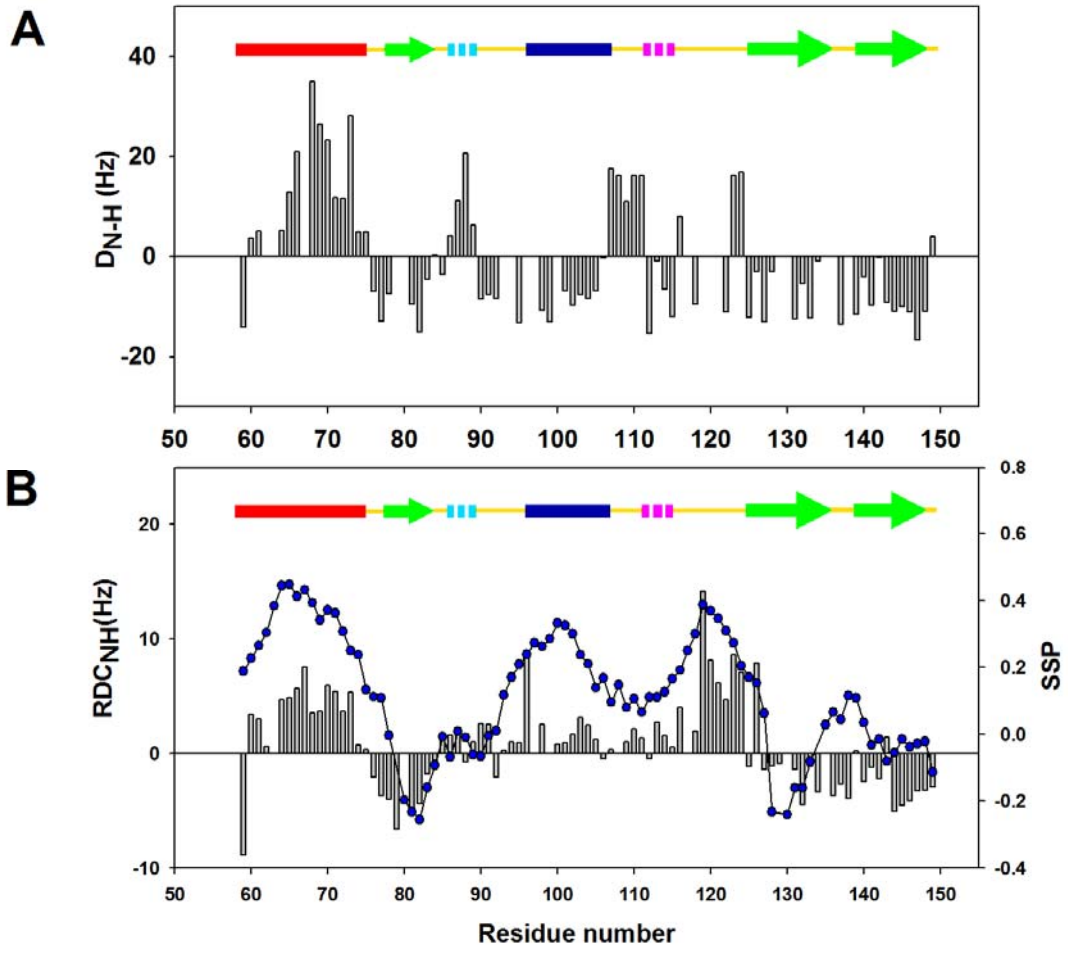


Figure 5-7. Residual dipolar couplings for CTL9 and the I98A mutant. Measurements were made at pH 5.7, 12 ^{o}C . (A) Plot of D_{NH} for the wild-type CTL9 *versus* residue numner. (B) Plot of D_{NH} (black bars) and SSP scores (●) for I98A CTL9 *versus* residue number. A schematic diagram of the elements of secondary structure of the native state of CTL9 is shown at the top of each panel.

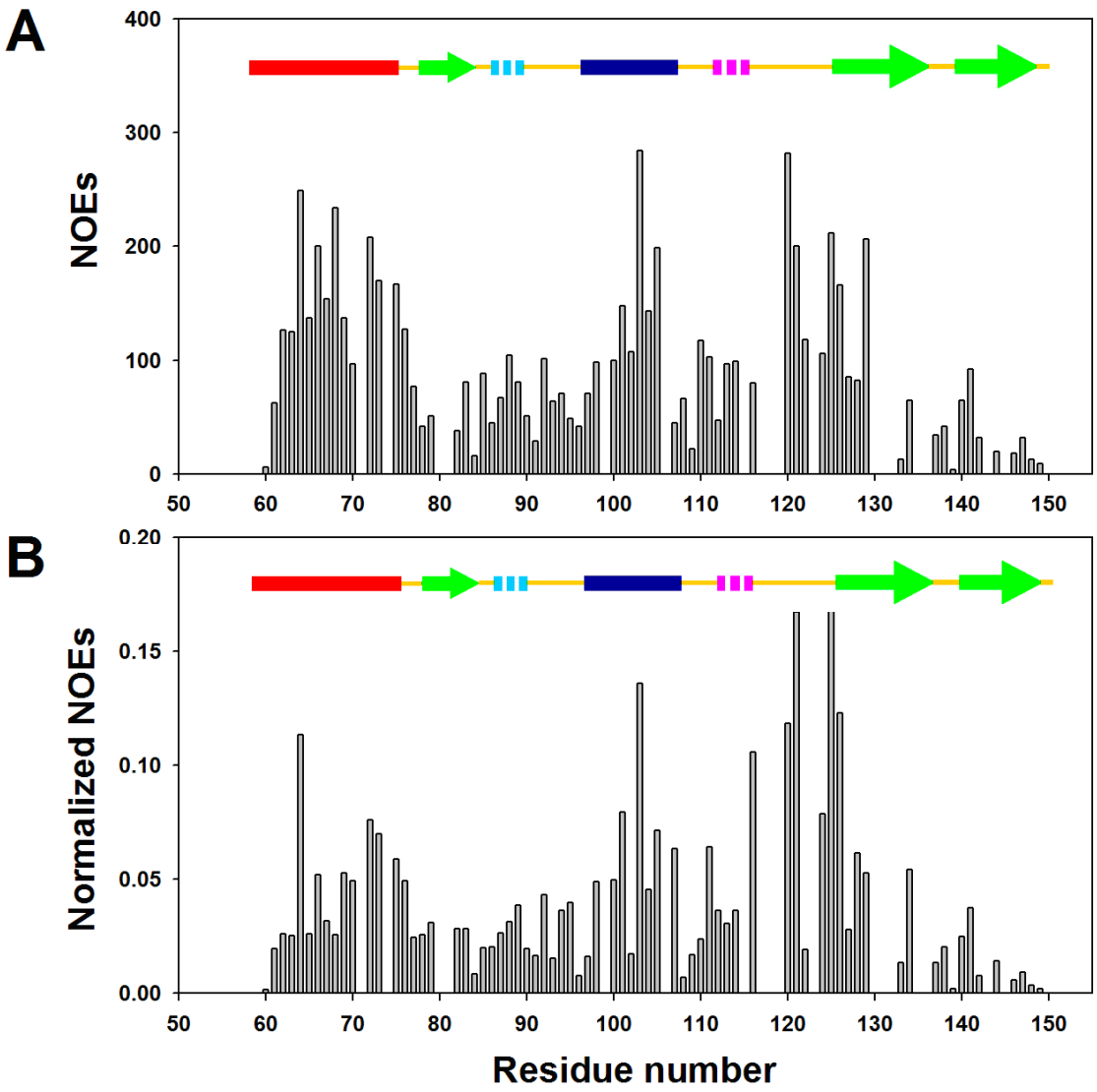


Figure 5-8. Summary of NOEs observed for I98A CTL9 at pH 5.7, 12 ^{o}C. (A) Plot of the volume of the sequential amide proton NOEs *versus* residue number. (B) Plot of normalized sequential amide proton NOEs *versus* residue numner. Peak volumes are normalized as the ratio of observed d_{NN}(i,i+1) NOE crosspeaks to the diagonal peaks.

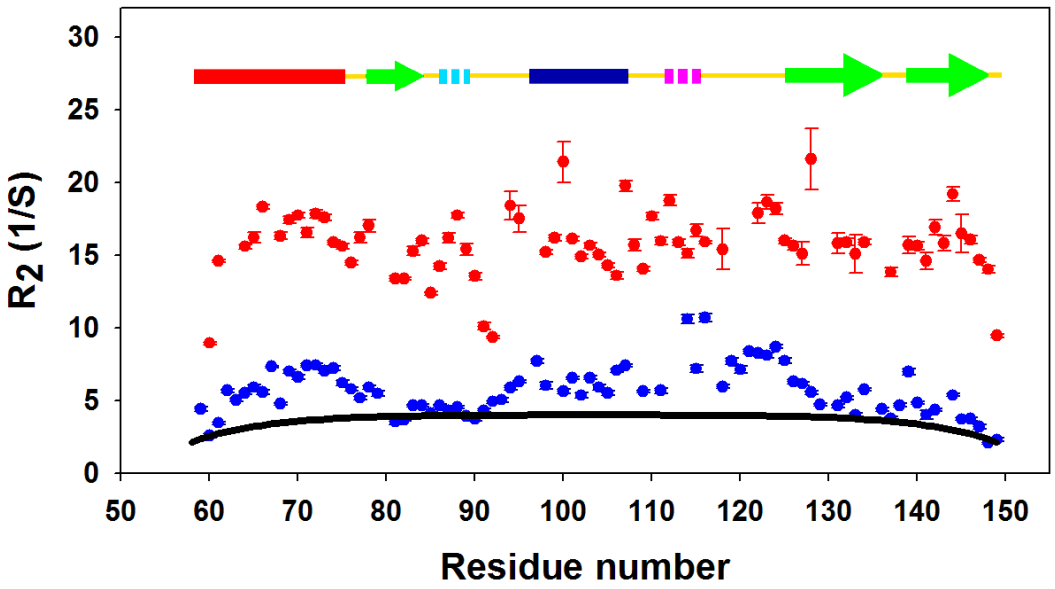


Figure 5-9. Plots of ^{15}N R_2 relaxation rates for wild-type CTL9 (●) and the cold denatured state of I98A CTL9 (●). The black solid line (-) is the best fit to the phenomenological model of Schwalbe and coworkers. Data was collected at 12^{o}C for both proteins.

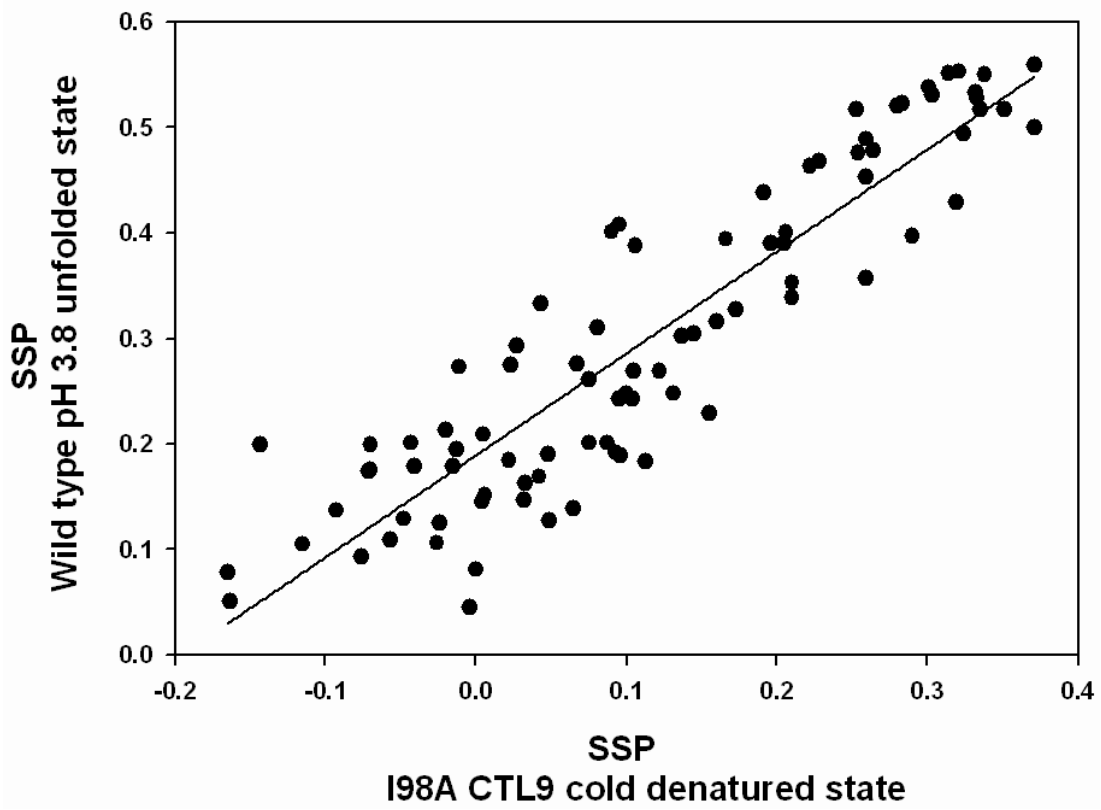


Figure 5-10. Comparison of the SSP scores determined for the pH 3.8 unfolded state of wild type CTL9 with those determined for the cold denatured state of I98A CTL9. The SSP scores for the pH 3.8 unfolded state were previously reported (*182*).

6. Correspondence between the overall dimensions and the amount of the secondary structure in the unfolded states of proteins

Measurement of molecular dimensions, usually taken as either the radius of hydration (R_h) or the radius of gyration (R_g), is of considerable importance in the structural characterization of native or partially folded states of proteins because it allows one to test the proposed correlations between the degree of compactness and the amounts of ordered secondary structural elements (*81, 214-215*). There appear to be a strong correlation between the degree of compactness and the CD detected relative amounts of secondary structure for native globular proteins and their partially folded intermediates. Uversky and Fink performed a statistical analysis of the secondary structure and the Stokes radii, R_s, for 41 different proteins in their native and partially folded states (*214*). They found that the data for both the native states and partially folded states can be described by an empirical linear equation: (correlation coefficient $r^2 = 0.97$):

$$\left(\frac{R_s^{\,U}}{R_s}\right)^3 = (1.047 \pm 0.010) * \frac{[\theta]_{222}}{[\theta]_{222}^{\,U}} - (0.31 \pm 0.12) \qquad (1)$$

Their study also suggests that all compact conformational states contain a high degree of well defined secondary structure and there are no non-compact well structured conformational states. In contrast, the computational studies by Pande and Rose indicate that global parameters such as R_g and R_h can be very insensitive to the presence of even significant amounts of secondary structure in unfolded proteins (*122, 125*). In the current chapter, the relationship between the secondary structure content and the radii of hydration of the various unfolded states of CTL9 determined in previous chapters is analyzed.

In the current study, the relative compactness can be evaluated as the decrease in the hydrodynamic volume of a given conformational state relative to that of the fully unfolded state, $(R_s^{\,U}/R_s)^3$, where the R_s is the Stokes radii and the superscript U indicates the fully unfolded conformation. The pH 2.5, 7.6 M urea unfolded state of CTL9 is taken as the fully unfolded state considering there is no detectable residual structure under such conditions. The content of ordered secondary structure can be determined by the CD ellipticity or the average SSP scores. Since negative SSP scores reflect the tendency to populate β-strand structure (*123*), the absolute values of the SSP scores of the pH 3.8 native state were used in the calculations. Based on CD detected pH titration experiments with CTL9, about 39% of the protein molecules populate the unfolded state and 61% populate the native state at pH 3.8 (*32*). The pH 3.8 native state adopts a very similar structure as the native state under neutral pH, thus the mean residue ellipticity of the pH 3.8 native state is taken as that of the native state at neutral pH, which is about -14,500 $deg.cm^{-1}.dmol^{-1}$. The observed CD ellipticity is the population weighted average of the signals of all possible conformation. As a result, the mean residue ellipticity of the pH 3.8 unfolded state was estimated to be about -3,450 $deg.cm^{-1}.dmol^{-1}$. The average SSP scores, CD ellipticity at 222 nm and the radii of hydration for various conformational states of CTL9 are summarized in Table 6-1.

The correlation deduced by Uversky and Fink is reproduced in Figure 6-1 which is adapted from their paper. The data for the various conformations of CTL9 are shown in the figure as red filled squares. Note that the CTL9 data was not included in the original study by Uversky and Fink. The data for the native state, the urea unfolded state and the pH 3.8 unfolded state follows the above correlation, however, a significant deviation is detected for the highly unfolded pH 2.0 state of CTL9.

The relative compactness and the relative average SSP scores for the different states of CTL9 appear to be correlated, with a correlation coefficient (r^2) of 0.78, but the limited data set (4 points) means that the statistical significance of the apparent moderate correlation is marginal as the p-value is 0.11 (Figure 6-2). Despite of the limited data points, the fact that the urea unfolded state and the pH 2.0 unfolded state have different amounts of residual secondary structure but the same R_h values argues that there is no strict correspondence between the compactness and the structure of a conformational state for unfolded proteins, thus supporting the computational studies by Pande and Rose. Nevertheless, such a correlation may hold true for native globular or partially folded proteins, as demonstrated by the CD data.

Residual structure in the unfolded states of proteins may play a critical role in the initiation of protein folding. The correlation between the hydrodynamic volume and the amounts of secondary structure are supportive of the hydrophobic collapse model of protein folding, which hypothesize that the hydrophobic collapse occurs concurrently with the formation of a compact conformation with significant amounts of secondary structure. It is not surprising that the correlation is abolished under highly denaturing conditions, in which the unfolded state is more favored.

Table 6-1. Average SSP scores, CD ellipticity at 222 nm and radii of hydration for various conformational states of CTL9

	Average SSP	$[\theta]_{222nm}$ (deg.cm^{-1}.dmol^{-1})	R_h (Å)
pH 2.0 unfolded	0.18	-6000	33.5
pH 3.8 unfolded	0.30	-3450[b]	28.1
pH 2.5 urea unfolded	0.03	-2500	33.6
pH 3.8 native	0.48[a]	-14500[c]	19.5

a. Calculated as the absolute value of SSP scores; b. estimated based on the ellipticity of the native state and and the relative population of the native and the unfolded states; c. estimated assuming the pH 3.8 native state has the same structure as the native state under neutral pH.

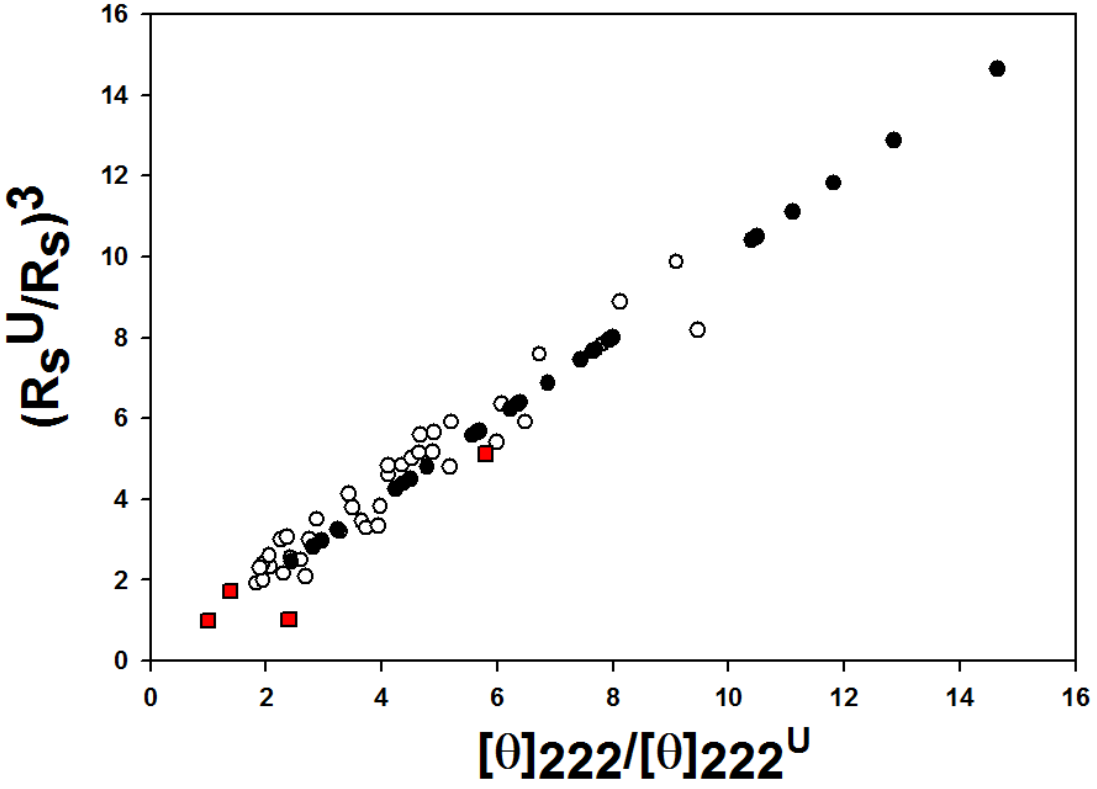


Figure 6-1. Correlation between the degree of compactness and the relative amounts of secondary structure calculated using CD signal as reported by Uversky and Fink (*214*). Filled and open circles represent the native globular proteins and their partially folded intermediates, respectively. The red filled squares correspond to the data for CTL9 obtained in previous chapters in this thesis, and were not included in the original analysis.

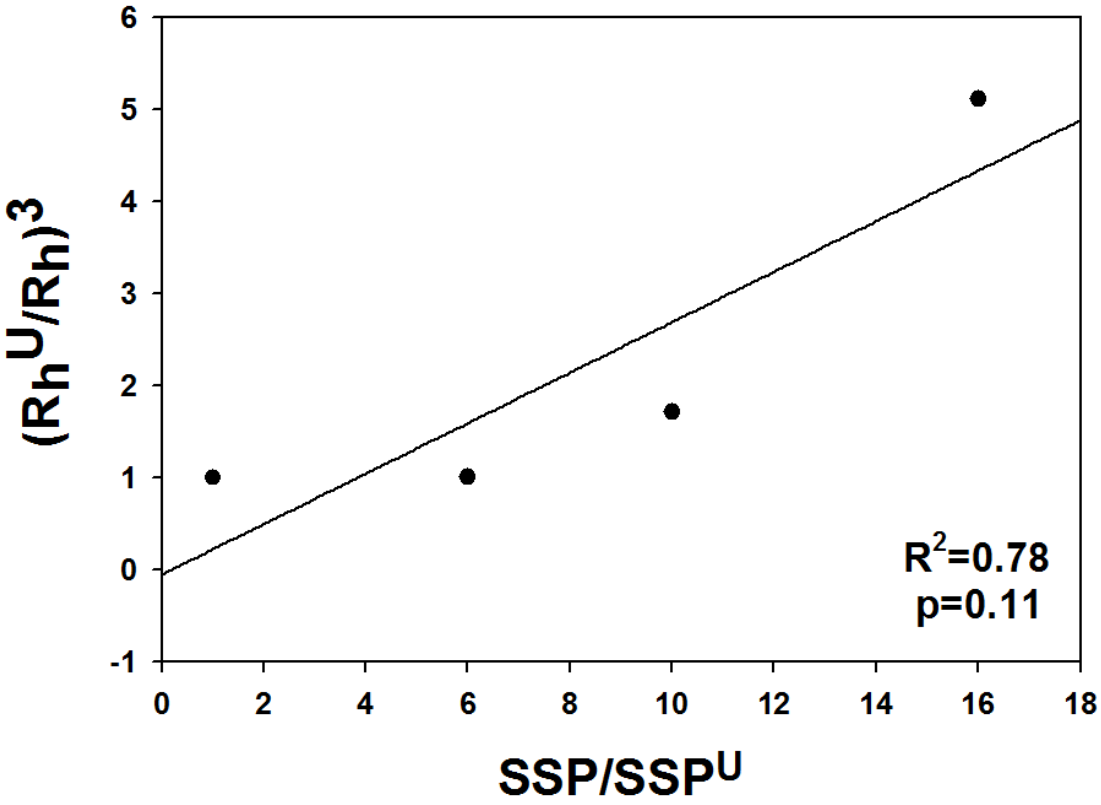


Figure 6-2. Correlation between the degree of compactness and the relative amounts of secondary structure calculated using SSP scores.

References

1. Anfinsen, C. B. (1973) Principles that govern folding of protein chains. *Science* 181, 223-230.
2. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* 65, 44-45.
3. Hart, T., Hosszu, L. L. P., Trevitt, C. R., Jackson, G. S., Waltho, J. P., Collinge, J., and Clarke, A. R. (2009) Folding kinetics of the human prion protein probed by temperature jump. *Proc. Natl. Acad. Sci. USA. 106*, 5651-5656.
4. Ensign, D. L., Kasson, P. M., and Pande, V. S. (2007) Heterogeneity even at the speed limit of folding: Large-scale molecular dynamics study of a fast-folding variant of the villin headpiece. *J. Mol. Biol. 374*, 806-816.
5. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., and Hofrichter, J. (2006) Sub-microsecond protein folding. *J. Mol. Biol. 359*, 546-553.
6. Kim, P. S., and Baldwin, R. L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* 51, 459-489.
7. Kim, P. S., and Baldwin, R. L. (1990) Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59, 631-660.
8. Dyson, H. J., and Wright, P. E. (1993) Peptide conformation and protein folding. *Curr. Opin. Struct. Biol.* 3, 60-65.
9. Karplus, M., and Weaver, D. L. (1994) Protein folding dynamics-the diffusion-collision model and experimental data. *Protein Sci.* 3, 650-668.
10. Roder, H., Elove, G. A., and Englander, S. W. (1988) Structural characterization of folding intermediates in cytochrome-C by H-exchange labelling and proton NMR. *Nature 335*, 700-704.
11. Kuroda, Y. (1993) Residual helical structure in the C-terminal fragment of cytochrome-C. *Biochemistry 32*, 1219-1224.
12. Dill, K. A. (1985) Theory for the folding and stability of globular-proteins. *Biochemistry* 24, 1501-1509.
13. Dill, K. A. (1990) The meaning of hydrophobicity. *Science* 250, 297-297.
14. Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry* 29, 7133-7155.
15. Dill, K. A., Bromberg, S., Yue, K. Z., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) Principles of protein folding-A perspective from simple exact models. *Protein Sci.* 4, 561-602.
16. Rackovsky, S., and Scheraga, H. A. (1977) Hydrophobicity, hydrophilicity and radial and orientational distributions of residues in native proteins. *Proc. Natl. Acad. Sci. USA*. 74, 5248-5251.
17. Fersht, A. R. (1995) Optimization of rates of protein folding-the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA*. 92, 10869-10873.
18. Fersht, A. R. (1997) Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* 7, 3-9.
19. Jackson, S. E., and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor-2. 1. Evidence for a 2-state transition. *Biochemistry* 30, 10428-10435.

20. Jackson, S. E., and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor-2. 2. Influence of proline isomerization on the folding kinetics and thermodynamics characterization of the transition state of folding. *Biochemistry* 30, 10436-10443.
21. Otzen, D. E., and Fersht, A. R. (1998) Folding of circular and permuted chymotrypsin inhibitor 2: Retention of the folding nucleus. *Biochemistry* 37, 8139-8146.
22. Nolting, B., Golbik, R., Neira, J. L., SolerGonzalez, A. S., Schreiber, G., and Fersht, A. R. (1997) The folding pathway of a protein at high resolution from microseconds to seconds. *Proc. Natl. Acad. Sci. USA*. 94, 826-830.
23. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes. *Science* 267, 1619-1620.
24. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways, and the energy landscape of protein folding - A synthesis. *Proteins* 21, 167-195.
25. Dobson, C. M., and Karplus, M. (1999) The fundemantals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* 9, 92-101.
26. Dobson, C. M., Sali, A., and Karplus, M. (1998) Protein folding: A perspective from theory and experiment. *Angew. Chem.-Int. Edit.* 37, 868-893.
27. Chan, H. S., and Dill, K. A. (1998) Protein folding in the landscape perspective: Chevron plots and non-Arrhenius kinetics. *Proteins: Struct. Funct. Genet.* 30, 2-33.
28. Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold. Des. 3*, R81-R91.
29. Kloss, E., Courtemanche, N., and Barrick, D. (2008) Repeat-protein folding: New insights into origins of cooperativity, stability, and topology. *Arch. Biochem. Biophys. 469*, 83-99.
30. Knott, M., and Chan, H. S. (2006) Criteria for downhill protein folding: Calorimetry, chevron plot, kinetic relaxation, and single-molecule radius of gyration in chain models with subdued degrees of cooperativity. *Proteins 65*, 373-391.
31. Chan, H. S., Shimizu, S., and Kaya, H. (2004) Cooperativity principles in protein folding. *Methods Enzymol.* 380*,* 350-379.
32. 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. 318*, 571-582.
33. Song, B. B., Cho, J. H., and Raleigh, D. P. (2007) Ionic-strength-dependent effects in protein folding: Analysis of rate equilibrium free-energy relationships and their lnterpretation. *Biochemistry 46*, 14206-14214.
34. Horng, J. C., Tracz, S. M., Lumb, K. J., and Raleigh, D. P. (2005) Slow folding of a three-helix protein via a compact intermediate. *Biochemistry 44*, 627-634.
35. McParland, V. J., Kalverda, A. P., Homans, S. W., and Radford, S. E. (2002) Structural properties of an amyloid precursor of beta(2)-microglobulin. *Nature Struct. Biol. 9*, 326-331.
36. Babu, C. R., Hilser, V. J., and Wand, A. J. (2004) Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. *Nat. Struct. Mol. Biol. 11*, 352-357.

37. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem. 24*, 1-95.
38. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266-280.
39. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m-values and heat capacity changes - relation to changes in accessible surface-areas of protein unfolding. *Protein Sci. 4*, 2138-2148.
40. Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., and Hofrichter, J. (1997) Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl. Acad. Sci. USA. 94*, 1779-1784.
41. Shastry, M. C. R., and Roder, H. (1998) Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol. 5*, 385-392.
42. Xu, Y., Oyola, R., and Gai, F. (2003) Infrared study of the stability and folding kinetics of a 15-residue beta-hairpin. *J. Am. Chem. Soc. 125*, 15388-15394.
43. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc. Natl. Acad. Sci. USA. 93*, 5759-5764.
44. Wang, M. H., Tang, Y. F., Sato, S. S., Vugmeyster, L., McKnight, C. J., and Raleigh, D. P. (2003) Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. *J. Am. Chem. Soc. 125*, 6032-6033.
45. Arora, P., Oas, T. G., and Myers, J. K. (2004) Fast and faster: A designed variant of the B-domain of protein A folds in 3 mu sec. *Protein Sci. 13*, 847-853.
46. Spector, S., and Raleigh, D. P. (1999) Submillisecond folding of the peripheral subunit-binding domain. *J. Mol. Biol. 293*, 763-768.
47. Korzhnev, D. M., and Kay, L. E. (2008) Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: An application to protein folding. *Accounts Chem. Res. 41*, 442-451.
48. Daggett, V., and Fersht, A. (2003) The present view of the mechanism of protein folding. *Nature Rev. Mol. Cell Biol. 4*, 497-502.
49. Dalby, P. A., Oliveberg, M., and Fersht, A. R. (1998) Folding intermediates of wild-type and mutants of barnase. I. Use of phi-value analysis and m-values to probe the cooperative nature of the folding pre-equilibrium. *J. Mol. Biol. 276*, 625-646.
50. 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.
51. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nature Rev. Mol. Cell Biol.* 6. 197-208
52. Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41, 415-427.
53. Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* 11, 739-756.
54. 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.

55. Wong, K. B., Freund, S. M. V., and Fersht, A. R. (1996) Cold denaturation of barstar: ^{1}H, ^{15}N and ^{13}C NMR assignment and characterisation of residual structure. *J. Mol. Biol.* 259, 805-818.
56. Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.* 469, 100-117.
57. Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation *in vitro*. *Biochemistry* 31, 8654-8660.
58. Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G., and Dobson, C. M. (1999) Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. USA. 96*, 3590-3594.
59. de Laureto, P. P., Taddei, N., Frare, E., Capanni, C., Costantini, S., Zurdo, J., Chiti, F., Dobson, C. M., and Fontana, A. (2003) Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol. 334*, 129-141.
60. Mishima, T., Ohkuri, T., Monji, A., Imoto, T., and Ueda, T. (2006) Amyloid formation in denatured single-mutant lysozymes where residual structures are modulated. *Protein Sci. 15*, 2448-2452.
61. Bussell, R., and Eliezer, D. (2001) Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. *J. Biol. Chem.* 276, 45996-46003.
62. Bussell, R., and Eliezer, D. (2004) Effects of Parkinson's disease-linked mutations on the structure of lipid-associated alpha-synuclein. *Biochemistry* 43, 4810-4818.
63. Platt, G. W., Victoria, J. M., Kalverda, A. P., Homans, S. W., and Radford, S. E. (2005) Dynamics in the Unfolded State of β2-microglobulin Studied by NMR. *J. Mol. Biol.* 346, 279-294.
64. Wirmer, J., Schlorb, C., Klein-Seetharaman, J., Hirano, R., Ueda, T., Imoto, T., and Schwalbe, H. (2004) Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew. Chem.-Int. Edit.* 43, 5780-5785.
65. Radhakrishnan, I., PerezAlvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator: Coactivator interactions. *Cell* 91, 741-752.
66. Radhakrishnan, I., Perez-Alvarado, G. C., Dyson, H. J., and Wright, P. E. (1998) Conformational preferences in the Ser(133)-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett.* 430, 317-322.
67. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I., and Wright, P. E. (1996) Structural studies of p21(Waf1/Cip1/Sdi1) in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. USA. 93*, 11504-11509.
68. Joerger, A. C., and Fersht, A. R. (2008) Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.* 77, 557-582.
69. Markus, M. A., Hinck, A. P., Huang, S. R., Draper, D. E., and Torchia, D. A. (1997) High resolution solution structure of ribosomal protein L11-C76, a helical protein with a flexible loop that becomes structured upon binding to RNA. *Nat. Struct. Biol.* 4, 70-77.

70. Wootton, J. C., and Federhen, S. (1993) Statistics of local complexity in amino-acid-sequences and sequence databases. *Comput. Chem. 17*, 149-163.
71. Wootton, J. C. (1994) Nonglobular domains in protein sequences-Automated segmentation using complexity measures. *Comput. Chem. 18,* 269-285.
72. Wootton, J. C., and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol. 266*, 554-571.
73. Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Bruce, J. E., Garner, E. C., Brown, C. J., Dunker, A. K., Smith, R. D., and Ackerman, E. J. (2001) Identification of intrinsic order and disorder in the DNA repair protein XPA. *Protein Sci.* 10, 560-571.
74. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22, 1302-1306.
75. Trovato, A., Seno, F., and Tosatto, S. C. E. (2007) The PASTA server for protein aggregation prediction. *Protein Eng. Des. Sel.* 20, 521-523.
76. Bryan, A. W., Menke, M., Cowen, L. J., Lindquist, S. L., and Berger, B. (2009) BETASCAN: Probable beta-amyloids identified by pairwise probabilistic analysis. *PLoS Comput. Biol.* 5, 11.
77. Ohgushi, M., and Wada, A. (1983) Molten-globule state - A compact form of globular proteins with mobile side-chains. *FEBS Lett.* 164, 21-24.
78. 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 of denaturant but is no more compact than the low-pH urea unfolded state. *Biochemistry* 47, 9565-9573.
79. 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.
80. Millet, I. S., Doniach, S., and Plaxco, K. W. (2002) Toward a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins. *Adv. Protein Chem. 349,* 241-262.
81. Jones, J. A., Wilkins, D. K., Smith, L. J., and Dobson, C. M. (1997) Characterisation of protein unfolding by NMR diffusion measurements. *J. Biomol. NMR* 10, 199-203.
82. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* 104, 3607-3622.
83. Eliezer, D. (2007) Characterization residual structure in disordered protein states using nuclear magnetic resonance. *Methods Mol. Biol.* 350, 49-67.
84. Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* 20, 71-75.
85. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207-234.
86. Wishart, D. S., 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.

87. 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. Biomol. NMR* 5, 67-81.
88. Crowhurst, K. A., and Forman-Kay, J. D. (2003) Aromatic and methyl NOES highlight hydrophobic clustering in the unfolded state of an SH3 domain. *Biochemistry* 42, 8687-8695.
89. 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.
90. Schwalbe, H., Fiebig, 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.
91. 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.
92. Ohnishi, S., and Shortle, D. (2003) Observation of residual dipolar couplings in short peptides. *Proteins* 50, 546-551.
93. Mohana-Borges, R., Goto, N. K., Kroon, G. J. A., Dyson, H. J., and Wright, P. E. (2004) Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.* 340, 1131-1142.
94. Hopkins, F. G. (1930) Denaturation of proteins by urea and related substances. *Nature* 126, 328-330.
95. Clark, J. H. (1945) The temperature coefficient of the urea denaturation of egg albumin. *J. Gen. Physiol.* 28, 539-545.
96. Prass, R. L., Isohashi, F., and Utter, M. F. (1980) Purification and characterization of an extra-mitochondrial acetyl coenzyme - A hydrolase from rat liver. *J. Biol. Chem.* 255, 5215-5223.
97. Vagelos, P. R., Alberts, A. W., and Martin, D. B. (1963) Studies on mechanism of activation of acetyl coenzyme. A carboxylase by citrate. *J. Biol. Chem.* 238, 533-540.
98. Privalov, P. L. (1990) Cold denaturation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 25, 281-305.
99. Chandler, D. (2005) Interfaces and the driving force of hydrophobic assembly. *Nature* 437, 640-647.
100. Privalov, P. L., and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* 39, 191-234.
101. Hatley, R. H. M., and Franks, F. (1986) Denaturation of lactate dehydrogenase at subzero temperatures. *Cryo-Lett.* 7, 226-233.
102. Hoffman, D. W., Davies, C., Gerchman, S. E., Kycia, J. H., Porter, S. J., White, S. W., and Ramakrishnan, V. (1994) Crystal structure of prokaryotic ribosomal protein L9 - A bilobed RNA binding protein. *Embo J.* 13, 205-212.
103. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) Ribosomal protein L9: A structure determination by the combined use of X-ray crystallography and NMR spectroscopy. *J. Mol. Biol*. 264, 1058-1071.

104. Sato, S., Kuhlman, B., Wu, W. J., and Raleigh, D. P. (1999) Folding of the multidomain ribosomal protein L9: The two domains fold independently with remarkably different rates. *Biochemistry* 38, 5643-5650.
105. Herr, A. J., Nelson, C. C., Wills, N. M., Gesteland, R. F., and Atkins, J. F. (2001) Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA. *J. Mol. Biol.* 309, 1029-1048.
106. Brimacombe, R., Gornicki, P., Greuer, B., Mitchell, P., Osswald, M., Rinkeappel, J., Schuler, D., and Stade, K. (1990) The three-dimensional structure and function of *escherichia coli* ribosomal RNA, as studied by cross-linking techniques. *Biochim. Biophys. Acta, 1050*, 8-13.
107. Walleczek, J., Redl, B., Stofflermeilicke, M., and Stoffler, G. (1989) Protein-protein cross-linking of the 50s ribosomal subunit of *escherichia coli* using 2-iminothiolane - identification of cross-links by immunoblotting techniques. *J. Biol. Chem.* 264, 4231-4237.
108. Li, Y., Shan, B., and Raleigh, D. P. (2007) The cold denatured state is compact but expands at low temperatures: Hydrodynamic properties of the cold denatured state of the C-terminal domain of L9. *J. Mol. Biol.* 368, 256-262.
109. Li, Y., Horng, J. C., and Raleigh, D. P. (2006) pH dependent thermodynamic and amide exchange studies of the C-terminal domain of the ribosomal protein L9: Implications for unfolded state structure. *Biochemistry* 45, 8499-8506.
110. Baldwin, R. L. (2002) A new perspective on unfolded proteins. *Advan. Protein Chem.* 361-367.
111. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J. Mol. Biol.* 338, 827-837.
112. Mok, Y. K., Elisseeva, E.L., Davidson, A. R. and Forman-Kay, J. D. (2001) Dramatic stabilization of an SH3 domain by a single substitution: roles of the folded and unfolded states. *J. Mol. Biol.* 307, 913-928.
113. 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.
114. Dill, K. A., and Shortle, D. (1991) Denatured states of proteins. *Annu. Rev. Biochem. 60*, 795-825.
115. 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.
116. Shortle, D. R. (1996) Structural analysis of non-native states of proteins by NMR methods. *Curr. Opin. Struct. Biol. 6*, 24-30.
117. Choy, W. Y., and Forman-Kay, J. D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* 308, 1011-1032.
118. Lindorff-Larson, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004) Determination of an ensemble of

structures representing the denatured state of the bovine acyl-coenzyme a binding protein. *J. Am. Chem. Soc.* 126, 3291-3299.
119. Zhang, O., and Forman-Kay, J. D. (1997) NMR studies of unfolded states of an SH3 domain and a stable mutant. *Biochemistry 44*, 15550-15560.
120. 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. USA* 101, 12491-12496.
121. Dyson, J. H., Wright, P. E. (2002) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Advan. Protein Chem. 62*, 311-340.
122. Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins. *Proc. Natl. Acad. Sci. USA 101*, 12497-12502.
123. Mittag, T., and Forman-Kay, J.D. (2007) Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol.* 17, 3-14.
124. Goldenberg, D. P. (2003) Computational Simulation of the Statistical Properties of Unfolded Proteins. *J. Mol. Biol. 326*, 1615-1633.
125. Zagrovic, B., and Pande, V.S. (2003) Structural correspondence between the α-helix and the random-flight chain resolves how the unfolded proteins can have native-like properties. *Nature Struct. Biol.* 10, 955-961.
126. Neri, D., Billeter,M., Wider, G., and Wuthrich, K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257, 1559-1563.
127. 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* 437, 1053-1056.
128. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., and Skelton, N.J. (1996) *Protein NMR spectroscopy: Principles and practice.* Academic Press, Inc.,San Diego, California.
129. 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.
130. 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.
131. 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.
132. Cantor, C. R., and Schimmel, R. (1980) *Biophysical Chemistry.* W.H. Freeman and Co. New York.
133. Kosen, P. A. (1989) Spin labeling of proteins. *Methods Enzymol.* 177, 86-121.
134. Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* 39, 5355-5365.

135. 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.
136. Johnson, B. A. (2004) Using NMRView to visualize and analyze NMR spectra of macromolecules. *Methods Mol. Biol.* 278, 313-352.
137. 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.
138. 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.
139. 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.
140. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol. 268*, 158-169.
141. Gillespie, J. R., and Shortle, D. (1997) Characterization 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.
142. Sung, Y. H., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol. 372*, 689-707.
143. 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.
144. Luo, P., and Baldwin, R. L. (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 36, 8413-8421.
145. Chin, D. H., Woody, R.W., Rohl, C. A., and Baldwin, R. L. (2002) Circular dichroism spectra of short, fixed-nucleus alanine helices. *Proc. Natl. Acad. Sci. USA 99*, 15416-15421.
146. 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.
147. 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.
148. 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.

149. 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.
150. 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.
151. 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.
152. 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.
153. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* 96, 432 - 440.
154. Kay, L.E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson. A,* 109, 129-133.
155. 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.
156. 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.
157. 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.
158. 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.
159. 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.
160. 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.
161. 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.
162. Bowler, B. E. (2007) Thermodynamics of protein denatured states. *Mol. Biosyst.* 3, 88-89.
163. 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.

164. 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.
165. 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.
166. 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.
167. Swintkruse, L., and Robertson, A. D. (1995) Hydrogen-bonds and the pH-dependence of ovomucoid 3rd domain stability. *Biochemistry* 34, 4724-4732.
168. 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.
169. 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.
170. 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.
171. 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.
172. 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.
173. Sanfelice, D., Tancredi, T., Politou, A., Pastore, A., and Temussi, P. A. (2009) Cold denaturation and aggregation: A comparative NMR study of Titin I28 in bulk and in a confined environment. *J. Am. Chem. Soc. 131*, 11662-11663.
174. Privalov, P. L., Griko, Y. V., Venyaminov, S. Y., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin. *J. Mol. Biol. 190*, 487-498.
175. Kunugi, S., and Tanaka, N. (2002) Cold denaturation of proteins under high pressure. *Biochim. Biophys. Acta 1595*, 329-344.
176. Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyarnal, S., and Akasaka, K. (2006) Cold denaturation of ubiquitin at high pressure. *Magn. Reson. Chem. 44*, S108-S113.
177. Jonas, J. (2002) High-resolution nuclear magnetic resonance studies of proteins. *Biochim. Biophys. Acta 1595*, 145-159.
178. Van Horn, W. D., Simorellis, A. K., and Flynn, P. F. (2005) Low-temperature studies of encapsulated proteins. *J. Am. Chem. Soc. 127*, 13553-13560.
179. Whitten, S. T., Kurtz, A. J., Pometun, M. S., Wand, A. J., and Hilser, V. J. (2006) Revealing the nature of the native state ensemble through cold denaturation. *Biochemistry 45*, 10163-10174.

180. Davidovic, M., Mattea, C., Qvist, J., and Halle, B. (2009) Protein cold denaturation as seen from the solvent. *J. Am. Chem. Soc.131*, 1025-1036.
181. R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
182. Shan, B., Eliezer, D., and Raleigh, D. P. (2009) The unfolded state of the C-terminal domain of the ribosomal protein L9 contains both native and non-Native Structure. *Biochemistry 48*, 4707-4719.
183. Segel, D. J., Fink, A. L., Hodgson, K. O., and Doniach, S. (1998) Protein denaturation: A small-angle X-ray scattering study of the ensemble of unfolded states of cytochrome c. *Biochemistry 37*, 12443-12451.
184. Snow, C. D., Nguyen, N., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics. *Nature 420*, 102-106.
185. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M., and Munoz, V. (2002) Experimental identification of downhill protein folding. *Science 298*, 2191-2195.
186. Simorellis, A. K., Van Horn, W. D., and Flynn, P. F. (2006) Dynamics of low temperature induced water shedding from AOT reverse micelles. *J. Am. Chem. Soc. 128*, 5082-5090.
187. Brandts, J.F. (1964) The thermodynamics of protein denaturation. I. Denaturation of chymotrypsinogen. *J. Am. Chem. Soc. 86*, 4291-4301.
188. Martin, S. R., Esposito, V., Rios, P. D. L., Pastore, A., and Temussi, P. A. (2008) Cold denaturation of yeast frataxin offers the clue to understand the effect of alcohols on protein stability. *J. Am. Chem. Soc. 130*, 9963-9970.
189. Lonhienne T., G. C., and Feller G. (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim. Biophys. Acta 1543*, 1-10.
190. Feller, G. and Gerday, C. (2003) Psychrophilic enzymes: Hot topics in cold denaturation. *Nat. Rev. Microbiol. 1*, 200-208.
191. Pometun, M. S., Peterson, R. W., Babu, C. R., and Wand, A. J. (2006) Cold denaturation of encapsulated ubiquitin. *J. Am. Chem. Soc. 128*, 10652-10653.
192. Van Horn, W. D., Ogilvie, M. E., and Flynn, P. F. (2009) Reverse micelle encapsulation as a model for intracellular crowding. *J. Am. Chem. Soc. 131*, 8030-8039.
193. Pastore, A., Martin, S. R., Politou, A., Kondapalli, K. C., Stemmler, T., and Temussi, P. A. (2007) Unbiased cold denaturation: Low- and high-temperature unfolding of yeast frataxin under physiological conditions. *J. Am. Chem. Soc. 129*, 5374-5375.
194. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson. 131*, 373-378.
195. Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science 278*, 1111-1114.

196. Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins-Information for structure determination in solution. *Proc. Natl. Acad. Sci. USA. 92*, 9279-9283.
197. Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. (2003) On the origin of residual dipolar couplings from denatured proteins. *J. Am. Chem. Soc. 125*, 15647-15650.
198. Ohnishi, S., Lee, A. L., Edgell, M. H., and Shortle, D. (2004) Direct demonstration of structural similarity between native and denatured eglin C. *Biochemistry 43*, 4064-4070.
199. Jha, A. K., Colubri, A., Freed, K. F., and Sosnick, T. R. (2005) Statistical coil model of the unfolded state: Resolving the reconciliation problem. *Proc. Natl. Acad. Sci. USA. 102*, 13099-13104.
200. Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R. W. H., and Blackledge, M. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA. 102*, 17002-17007.
201. Wuthrich, K. (1986) *NMR of proteins and nucleic acids.* John Wiley & Sons, Inc., New York.
202. Clore, G. M., and Gronenborn, A. M. (1983) Theory of the time-dependent transferred nuclear overhauser effect-Applications to structural-analysis of ligand protein complexes in solution. *J. Magn. Reson. 53*, 423-442.
203. Post, C. B. (2003) Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Curr. Opin. Struct. Biol. 13*, 581-588.
204. Bodner, C. R., Dobson, C. M., and Bax, A. (2009) Multiple Tight Phospholipid-Binding Modes of alpha-Synuclein Revealed by Solution NMR Spectroscopy. *J. Mol. Biol. 390*, 775-790.
205. Mok, Y. K., Kay, C. M., Kay, L. E., and Forman-Kay, J. (1999) NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol. 289*, 619-638.
206. 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.
207. Palmer, A. G. (1997) Probing molecular motion by NMR. *Curr. Opin. Struct. Biol. 7*, 732-737.
208. Kay, L. E. (1998) Protein dynamics from NMR. *Nat. Struct. Biol. 5*, 513-517.
209. Antonino, L. C., Kautz, R. A., Nakano, T., Fox, R. O., and Fink, A. L. (1991) Cold denaturation and (H2O)-H-2 stabilization of a staphylococcal nuclease mutant. *Proc. Natl. Acad. Sci. USA. 88*, 7715-7718.
210. Szyperski, T., Mills, J. L., Perl, D., and Balbach, J. (2006) Combined NMR-observation of cold denaturation in supercooled water and heat denaturation enables accurate measurement of Delta C-p of protein unfolding. *Eur. Biophys. J. Biophys. Lett. 35*, 363-366.
211. Jonas, J., and Jonas, A. (1994) High-pressure NMR-spectroscopy of proteins and membranes. *Annu. Rev. Biophys. Biomol. Struct. 23*, 287-318.
212. Jonas, J. (1997) Cold denaturation of proteins. *ACS Symp. S. 676*, 310-323.

213. Jonas, J., Ballard, L., and Nash, D. (1998) High-resolution, high-pressure NMR studies of proteins. *Biophys. J. 75*, 445-452.
214. Uversky, V. N., and Fink, A. L. (2002) The chicken-egg scenario of protein folding revisited. *FEBS Lett. 515*, 79-83.
215. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry 38*, 16424-16431.

Appendix 1. CTL9 backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments at pH 2.0

Residue	NH	N	CA	CB	CO	HA	HB	HB'
A58			51.92	19.18	173.49	4.10	1.56	--
A59	8.61	123.73	52.68	19.15	177.65	4.36	1.41	--
E60	8.46	120.28	55.99	28.73	176.06	4.36	2.05	--
E61	8.43	121.88	55.99	28.56	176.18	4.38	2.06	--
L62	8.28	123.78	55.73	42.53	177.65	4.31	1.64	--
A63	8.22	123.83	53.20	18.91	178.18	4.24	1.41	--
N64	8.30	117.35	53.70	38.62	175.76	4.62	2.83	--
A65	8.12	123.98	53.70	18.93	178.64	4.21	1.44	--
K66	8.13	119.44	57.50	32.62	177.35	4.18	1.82	--
K67	8.03	121.32	57.00	32.63	177.30	4.24	1.82	--
L68	8.08	122.32	55.99	42.27	177.94	4.29	1.63	--
K69	8.15	121.26	57.28	32.62	177.29	4.21	1.84	--
E70	8.17	120.22	56.75	28.57	176.65	4.27	2.08	--
Q71	8.30	121.01	56.74	29.08	176.51	4.29	2.09	--
L72	8.18	122.61	55.74	42.27	177.88	4.30	1.66	--
E73	8.19	120.48	56.23	28.57	176.12	4.34	2.07	--
K74	8.16	121.95	56.62	32.89	176.46	4.30	1.80	--
L75	8.17	123.20	55.35	42.53	177.53	4.43	1.64	--
T76	8.16	116.09	62.09	69.95	174.31	4.73	4.31	--
V77	8.15	123.11	62.33	32.89	176.01	4.23	2.08	--
T78	8.30	119.64	61.82	69.95	174.07	4.36	4.10	--
I79	8.29	125.96	58.53	38.72	174.45	--	--	--
P80	--	--	63.35	32.12	176.61	4.38	2.30	1.90
A81	8.32	124.48	52.54	19.17	177.94	4.28	1.39	--
K82	8.21	120.50	56.23	33.12	176.36	4.29	1.78	--
A83	8.26	125.38	52.69	19.18	178.29	4.30	1.39	--
G84	8.38	108.31	45.33		174.42	3.97	--	--
E85	8.24	119.73	55.98	28.71	176.82	4.40	2.17	2.01
G86	8.53	110.01	45.58		174.77	3.97	--	--
G87	8.25	108.62	45.32		174.25	3.96	--	--
R88	8.09	120.39	56.36	30.60	176.18	4.30	1.76	--
L89	8.22	122.93	55.23	42.28	177.12	4.30	1.51	--
F90	8.22	120.81	58.02	39.74	176.27	4.60	3.17	3.03
G91	8.30	110.33	45.24		174.07	3.94	--	--
S92	8.16	115.70	58.27	64.11	174.74	4.49	3.89	--
I93	8.25	122.54	61.58	38.72	176.70	4.29	1.93	--
T94	8.18	117.44	62.08	69.95	174.95	4.34	--	--
S95	8.31	117.95	59.28	63.71	175.19	4.42	3.92	--
K96	8.27	123.33	57.25	32.88	176.82	4.28	1.80	--
Q97	8.21	120.88	56.87	29.70	177.06	4.26	2.07	--
I98	8.12	122.24	62.33	38.47	176.83	4.05	1.87	--
A99	8.26	126.44	53.45	18.91	178.76	4.26	1.43	--
E100	8.27	119.09	56.85	28.57	177.09	4.29	2.09	--
S101	8.23	116.61	59.29	63.60	175.36	4.40	3.95	--
L102	8.17	123.55	56.39	42.14	178.35	4.28	1.69	1.60
Q103	8.13	119.81	56.76	28.83	176.65	4.23	2.08	--
A104	8.10	123.90	53.18	18.92	178.29	4.06	1.41	--
Q105	8.15	118.42	56.49	29.07	176.42	4.22	2.05	--

H106	8.38	118.49	55.58	28.82	175.07	4.70	3.35	3.21
G107	8.38	109.38	45.58		174.13	3.96	--	--
L108	8.04	121.77	55.48	42.53	177.59	4.34	1.61	--
K109	8.32	122.31	56.75	32.63	176.65	4.34	1.80	--
L110	8.17	123.22	55.47	42.53	177.18	4.33	1.60	--
D111	8.41	120.16	53.19	38.21	175.36	4.67	2.91	--
K112	8.26	122.21	57.00	32.88	176.83	4.25	1.80	--
R113	8.18	121.47	56.75	30.59	176.60	4.26	1.82	--
K114	8.22	122.36	57.01	32.89	177.06	4.24	1.80	--
I115	8.08	122.55	61.58	38.62	176.42	4.08	1.84	--
E116	8.33	124.00	55.99	28.56	176.59	4.37	2.06	--
L117	8.26	123.80	55.48	42.28	177.62	4.33	1.63	--
A118	8.29	124.30		19.18	178.35	4.24	1.41	--
D119	8.30	117.34	53.45	37.95	175.42	4.62	2.91	--
A120	8.06	124.27	53.19	18.92	178.09	4.30	1.41	--
I121	7.95	119.39	61.82	38.46	176.83	4.06	1.87	--
R122	8.18	124.50	56.50	30.60	176.24	4.29	1.80	--
A123	8.17	124.78	52.68	19.04	177.82	4.29	1.39	--
L124	8.10	121.14	55.48	42.53	178.06	4.29	1.65	1.54
G125	8.24	108.69	45.32		173.90	3.92	--	--
Y126	7.95	119.82	58.02	38.73	175.95	4.63	3.03	--
T127	8.03	115.91	61.57	69.95	173.66	4.32	4.14	--
N128	8.36	121.55	53.19	38.98	174.60	4.70	2.76	--
V129	8.04	122.19	60.06	32.64	174.36	--	2.02	--
P130	--	--	63.35	32.03	177.06	4.44	2.29	1.85
V131	8.17	120.98	62.58	32.88	176.18	4.04	2.03	--
K132	8.33	125.39	55.99	32.89	176.12	4.32	1.73	--
L133	8.21	123.71	55.22	42.54	176.83	4.28	1.58	1.44
H134	8.50	119.29			172.20	--	3.19	3.11
P135	--	--	63.35	32.11	176.82	4.44	2.29	1.92
E136	8.53	121.21	55.74	28.82	176.09	4.43	2.06	--
V137	8.29	122.32	62.34	32.87	176.30	4.20	2.09	--
T138	8.24	118.82	62.08	69.95	174.25	4.32	4.19	--
A139	8.36	127.04	52.68	19.17	177.77	4.39	1.42	--
T140	8.12	114.08	62.08	69.94	174.48	4.28	4.17	--
L141	8.18	125.11	55.36	42.53	177.12	4.37	1.61	--
K142	8.31	122.98	56.24	32.88	176.30	4.35	1.75	--
V143	8.00	120.99	62.33	32.88	175.83	4.08	2.00	--
H144	8.65	122.86	55.22	28.82	174.13	4.76	3.19	--
V145	8.28	122.98	62.33	32.89	176.10	4.18	2.06	--
T146	8.33	119.12	62.08	69.94	174.25	4.34	4.16	--
E147	8.41	123.41	55.62	28.82	175.65	4.42	2.11	2.00
Q148	8.45	122.53	55.99	29.58	175.77	4.36	2.05	
K149	8.49	123.97	55.73	32.63	178.41		1.76	

Appendix 2. CTL9 backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments at pH 2.5, in 7.6 M urea

Residue	NH	N	CA	CB	CO	HA	HB	HB'
A58			51.88	19.08	173.46	4.07	1.49	
A59	8.58	123.90	52.47	19.06	177.35	4.32	1.34	
E60	8.40	120.23	55.99	29.02	175.87	4.31	1.96	
E61	8.44	122.19	55.98	29.02	175.78			
L62								
A63			52.48	19.07	177.52	4.23	1.31	
N64	8.31	117.78	53.06	38.99	175.01	4.64	2.75	
A65	8.12	124.43	53.06	19.06	177.50	4.23	1.32	
K66	8.21	120.45	56.58	33.14	176.56	4.20	1.73	
K67	8.24	122.76	56.57	33.14	176.48			
L68			54.83	42.51	177.35	4.29	1.51	
K69	8.43	122.77	56.59	33.13	176.53	4.23	1.71	
E70	8.35	121.89	56.02	29.02	175.87	4.29	1.96	
Q71	8.49	122.32	55.98	29.61	175.76			
L72			55.42	42.22	177.59	4.26	1.55	
E73								
K74								
L75			54.82	42.51	177.37	4.41	1.54	
T76	8.27	116.53	61.87	70.06	174.27	4.32	4.12	
V77	8.19	122.81	61.86	33.13	175.96	4.18	2.00	
T78	8.26	119.04	61.86	70.06	174.04	4.33	4.05	
I79	8.26	125.23	58.35	39.00	174.31			
P80			63.03	31.96	176.44	4.32	2.21	1.83
A81	8.30	124.61	53.04	19.07	177.69	4.23	1.32	
K82	8.23	120.79	56.58	33.14	176.27			
A83	8.31	125.53	52.50	19.06	178.06	4.27	1.34	
G84	8.32	108.33	45.44		174.17	3.94		
E85	8.20	119.58	56.00	29.02	176.54	4.35	2.10	1.94
G86	8.45	109.90	45.45		174.43	3.94		
G87	8.20	108.47	45.43		173.93	3.93		
R88	8.14	120.53	56.27	30.80	176.03	4.25	1.66	
L89	8.25	123.37	54.82	42.52	176.93	4.26	1.42	
F90	8.31	121.17	57.76	39.60	176.15	4.59	3.02	
G91	8.30	110.31	45.43		173.80	3.92		
S92	8.17	115.67	58.34	64.20	174.60	4.48	3.79	
I93	8.27	122.58	61.27	38.85	176.41	4.23	1.83	
T94	8.17	117.51	61.84	70.06	174.60	4.25	4.23	
S95	8.26	118.16	58.43	64.18	174.60	4.39	3.82	
K96	8.31	123.42	56.57	33.13	176.43	4.24	1.71	
Q97	8.30	121.56	56.58	29.61	176.16	4.26	1.95	
I98	8.16	122.62	61.27	38.97	176.10	4.07	1.78	
A99	8.31	127.69	53.05	19.05	177.76	4.26	1.33	
E100	8.27	119.87	56.00	29.03	176.28	4.29	1.99	
S101	8.30	116.86	58.34	63.84	174.77			
L102								
Q103								
A104								
Q105	8.23	119.24	55.99	29.61	176.00	4.23	1.98	

H106	8.48	119.09	55.39	29.02	174.73	4.70	3.18	
G107	8.40	109.79	45.42		173.66	3.95		
L108	8.10	121.87	55.10	42.52	177.21	4.29	1.49	
K109	8.46	123.36	56.24	33.11	176.20			
L110			54.82	42.50	176.94	4.30	1.42	
D111	8.56	120.91	53.06	38.39	175.03	4.25	2.81	
K112	8.32	122.67	56.57	33.13	176.29	4.68	1.70	
R113	8.32	122.67	55.99	29.60	176.16	4.24	1.73	
K114	8.42	123.87	56.58	32.56	176.27	4.25	1.66	
I115	8.25	123.37	61.25	39.00	175.93	4.23	1.72	
E116								
L117								
A118			52.45	19.06	177.55	4.23	1.32	
D119	8.33	117.99	53.05	38.39	174.78			
A120			52.45	19.06	177.43	4.26	1.31	
I121	8.00	120.10	61.25	38.97	176.19	4.08	1.73	
R122								
A123								
L124	8.19	121.84	55.39	42.50	177.74	4.25	1.51	
G125	8.26	109.05	45.42		173.59	3.85		
Y126	8.02	119.92	57.75	39.00	175.94	4.60	2.93	
T127	8.09	115.73	61.83	70.06	173.74	4.28	4.09	
N128	8.38	121.59	53.06	38.99	174.59	4.68	2.71	
V129	8.02	121.46	60.01	32.38	174.23			
P130			63.02	31.96	176.66	4.39	2.19	1.77
V131	8.13	120.98	62.74	32.53	176.01	4.02	1.92	
K132	8.32	125.60	55.99	33.13	175.70			
L133			54.82	42.51	177.01	4.23	1.33	
H134	8.64	119.69	53.06	28.43	172.34			
P135			63.01	31.98	176.61	4.39	2.23	1.86
E136	8.53	121.46	55.99	29.03	175.92	4.39	1.96	
V137	8.27	122.07	62.44	33.11	176.20	4.16	1.92	
T138	8.22	118.41	61.53	70.06	174.18	4.30	4.16	
A139	8.28	126.75	52.49	19.43	177.57	4.37	1.34	
T140	8.10	113.99	61.85	70.06	174.40	4.16	4.14	
L141	8.20	125.15	54.84	42.52	177.00	4.27	1.50	
K142	8.39	123.43	55.99	33.12	176.28	4.31	1.68	
V143	8.09	121.30	62.42	33.12	175.79	4.05	1.93	
H144	8.67	123.03	54.85	29.03	174.11	4.73	3.13	
V145	8.31	123.19	62.43	33.13	175.98	4.18	2.00	
T146	8.31	118.92	61.84	69.80	174.27	4.27	4.25	
E147	8.38	123.30	55.99	29.01	176.16	4.35	1.95	
Q148	8.46	122.51	55.99	29.61	175.69	4.29	1.97	
K149	8.46	123.73	55.98	32.55	178.57			

Appendix 3. ^{15}N R2 relaxation rates for the pH 2.0 and pH 2.5, 7.6 M urea unfolded states of CTL9 and calculated R2 rates using random coil model

Residue	R_2 urea (sec^{-1})	R_2 pH 2.0 (sec^{-1})	Model values (sec^{-1})
A58			1.6
A59	1.1± 0.1	0.93 ± 0.02	1.8
E60	1.9 ± 0.1	1.7 ± 0.02	1.9
E61		1.1 ± 0.02	2.0
L62		1.8 ± 0.03	2.2
A63		4.4 ± 0.04	2.3
N64	2.5 ± 0.03	1.9 ± 0.02	2.3
A65	2.1± 0.2	2.8 ± 0.02	2.4
K66	3.2 ± 0.03	2.6 ± 0.03	2.5
K67	3.5 ± 0.04	3.1 ± 0.03	2.5
L68		3.7 ± 0.04	2.6
K69	3.0 ± 0.03	3.1± 0.03	2.6
E70	4.0 ± 0.04	3.6 ± 0.04	2.7
Q71	3.7 ± 0.04	3.3 ± 0.03	2.7
L72		3.6 ± 0.04	2.8
E73	3.2 ± 0.03		2.8
K74		3.1± 0.04	2.8
L75		3.2 ± 0.03	2.8
T76	3.4 ± 0.02	2.6 ± 0.03	2.9
V77	3.2 ± 0.02	3.1 ± 0.02	2.9
T78	3.3 ± 0.01	3.6 ± 0.04	2.9
I79	4.1 ± 0.1	3.8 ± 0.03	2.9
P80			2.9
A81		2.0 ± 0.03	2.9
K82			2.9
A83		2.3 ± 0.03	2.9
G84	3.0 ± 0.04	2.7 ± 0.03	2.9
E85	3.0 ± 0.04	2.1 ± 0.03	3.0
G86	2.8 ± 0.1	1.6 ± 0.03	3.0
G87	2.5 ± 0.05	1.8 ± 0.02	3.0
R88	2.8 ± 0.03	1.5 ± 0.02	3.0
L89		2.2 ± 0.03	3.0
F90	2.5 ± 0.04	3.1 ± 0.02	3.0
G91	2.5 ± 0.05	2.2 ± 0.02	3.0
S92	2.4 ± 0.02	2.3 ± 0.02	3.0
I93	3.7 ± 0.04		3.0
T94	4.4 ± 0.05	3.0 ± 0.03	3.0
S95		4.2 ± 0.03	3.0
K96		3.8 ± 0.03	3.0
Q97	3.2 ± 0.05		3.0
I98	4.3 ± 0.06	4.4 ± 0.04	3.0
A99	4.4 ± 0.1	4.7 ± 0.04	3.0
E100	3.5 ± 0.02	4.7 ± 0.04	3.0
S101	4.3 ± 0.04	4.0 ± 0.03	3.0
L102		1.9 ± 0.03	3.0
Q103		3.5 ± 0.03	3.0
A104	3.4 ± 0.04		3.0

Q105	3.5 ± 0.03	2.6 ± 0.03	3.0
H106	3.4 ± 0.03	2.8 ± 0.03	3.0
G107	3.4 ± 0.1	3.7 ± 0.03	3.0
L108	4.2 ± 0.04	2.9 ± 0.03	3.0
K109	3.4 ± 0.03		3.0
L110			3.0
D111	3.4 ± 0.02	2.9 ± 0.03	3.0
K112		3.0 ± 0.03	3.0
R113		5.0 ± 0.05	3.0
K114	4.7 ± 0.03	3.0 ± 0.03	3.0
I115		4.2 ± 0.03	3.0
E116		4.9 ± 0.04	3.0
L117			3.0
A118		1.7 ± 0.02	3.0
D119	3.0 ± 0.02		3.0
A120		3.8 ± 0.03	3.0
I121	3.4 ± 0.03	3.5 ± 0.03	3.0
R122	4.3 ± 0.05	3.3 ± 0.03	3.0
A123		3.2 ± 0.03	2.9
L124	2.9 ± 0.05	3.2 ± 0.03	2.9
G125	3.6 ± 0.05	2.6 ± 0.03	2.9
Y126	3.8 ± 0.05	3.2 ± 0.04	2.9
T127	3.8 ± 0.05	2.0 ± 0.03	2.9
N128	3.9 ± 0.04	2.8 ± 0.03	2.9
V129	3.6 ± 0.03	2.3 ± 0.04	2.9
P130			2.9
V131	3.5 ± 0.03	2.9 ± 0.04	2.9
K132		3.1 ± 0.03	2.8
L133		3.2 ± 0.04	2.8
H134	4.2 ± 0.02	2.0 ± 0.03	2.8
P135			2.8
E136	3.3 ± 0.03	2.2 ± 0.03	2.7
V137	3.7 ± 0.1	3.1 ± 0.04	2.7
T138	3.8 ± 0.3	2.5 ± 0.03	2.6
A139	4.2 ± 0.1	2.3 ± 0.03	2.6
T140	3.6 ± 0.1	2.2 ± 0.02	2.5
L141		2.6 ± 0.03	2.5
K142		1.6 ± 0.02	2.4
V143	2.5 ± 0.03	2.0 ± 0.03	2.3
H144	3.3 ± 0.06	1.8 ± 0.04	2.3
V145	3.0 ± 0.05	1.5 ± 0.02	2.1
T146	3.3 ± 0.05	2.4 ± 0.03	2.0
E147		1.3 ± 0.02	1.9
Q148	1.7 ± 0.03	1.8 ± 0.02	1.8
K149	2.2 ± 0.05	1.4 ± 0.02	1.6

Appendix 4. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-1H crosspeaks in the HSQC spectrum of K74C, K109C and D119C mutants of CTL9 at pH 2.0

Residue	K74C	K109C	D119C
A59	0.72	1.04	1.18
E60	0.53	1.01	1.03
E61	0.44	0.81	0.91
L62		0.75	0.88
A63	0.32	0.21	0.80
N64	0.65	0.64	0.65
A65			0.92
K66	0.48	1.00	0.92
K67		--	0.92
L68		--	0.94
K69	0.27		0.65
E70	0.21	0.81	0.92
Q71		0.82	0.86
L72		0.69	0.83
E73			0.77
K74			0.75
L75		0.57	
T76		0.60	0.81
V77			0.64
T78		0.44	0.59
I79	0.21	0.50	0.64
P80			
A81	0.23	0.33	0.53
K82	0.20	0.61	0.66
A83	0.37	0.76	0.72
G84		0.53	0.63
E85	0.44	0.72	0.76
G86	0.38	0.59	0.68
G87			0.84
R88	0.35	0.60	0.60
L89	0.28	0.48	0.34
F90	0.23	0.35	0.41
G91	0.17	0.29	0.30
S92	0.24	0.47	0.53
I93	0.30	0.35	0.30
T94	0.31	0.47	0.40
S95	0.29	0.37	0.34
K96	0.43		0.55
Q97	0.21	0.29	0.43
I98			0.40
A99		0.26	0.11
E100			0.00
S101	0.26	0.35	0.00
L102	0.64		0.31
Q103	0.43		0.36
A104	0.40	0.30	0.37
Q105	0.30	0.28	0.32

H106	0.62	0.21	0.64
G107	0.57	0.21	0.48
L108	0.36	0.00	0.33
K109	0.39		0.42
L110			
D111	0.27		0.23
K112	0.40		0.41
R113	0.48		0.20
K114			
I115		0.71	0.21
E116	0.37	0.54	0.00
L117			
A118	0.44	0.73	0.00
D119	0.38	0.33	0.00
A120	0.32	0.30	
I121	0.30	0.22	0.00
R122	0.39		0.00
A123			0.00
L124		0.21	
G125	0.27	0.82	0.00
Y126	0.41	0.44	0.24
T127	0.38	0.45	0.19
N128	0.45	0.45	0.28
V129	0.57	0.74	0.58
P130			
V131		0.48	0.39
K132	0.44	0.56	0.37
L133	0.40	0.68	0.43
H134	0.52	0.63	0.36
P135			
E136	0.44	0.55	0.32
V137	0.80	0.84	0.85
T138	0.82	0.82	0.92
A139	0.71	0.82	0.78
T140	0.74	0.94	0.91
L141	0.76	0.74	--
K142	0.82	0.91	1.00
V143	0.88	0.79	0.94
H144	0.79	0.96	0.89
V145			
T146	0.92	0.91	0.94
E147	0.93	1.01	1.04
Q148	0.84	0.99	1.01
K149	0.85	1.15	0.96

Appendix 5. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-1H crosspeaks in the HSQC spectrum of K74C, K109C and D119C mutants of CTL9 at pH 2.0 in 8M urea

Residue	K74C	K109C	D119C
A59	1.23	1.00	1.05
E60	0.84	1.03	0.99
E61	0.81	1.13	1.01
L62			
A63			1.03
N64	0.77	0.94	1.02
A65	0.94	1.04	1.00
K66	0.79	0.98	1.19
K67		1.09	1.09
L68			
K69	0.42	0.97	0.98
E70	0.26	0.99	1.08
Q71		0.97	1.11
L72			
E73		0.37	1.09
K74			
L75			
T76		0.98	1.12
V77		1.10	1.17
T78		0.85	1.14
I79		0.90	1.06
P80			
A81			0.99
K82	0.89	0.94	0.98
A83			
G84	0.75	1.04	1.04
E85	0.89	1.09	1.11
G86	0.83	1.05	0.96
G87	0.74	0.94	1.05
R88	0.92	0.96	1.12
L89		0.84	0.79
F90	0.76	0.96	0.98
G91	0.58	0.87	0.95
S92	0.64	0.84	1.06
I93	0.76	0.96	1.06
T94	0.79	0.97	0.95
S95	0.83	0.83	--
K96	0.85	0.83	1.04
Q97	0.85	0.79	1.09
I98	0.48		1.04
A99	0.77	0.50	0.93
E100	0.79	0.43	0.73
S101		0.48	0.93
L102			
Q103			
A104	0.91		--
Q105		0.34	0.79

H106	0.89	0.26	1.06
G107	0.88		0.83
L108	1.03		0.85
K109	1.05		0.95
L110			1.12
D111	0.87		0.75
K112	0.91		
R113		0.32	0.86
K114	0.89		0.53
I115	0.81		
E116	0.80		
L117			
A118			
D119	0.95		
A120			
I121	0.92		
R122	0.91	0.37	
A123		0.88	
L124	0.98	0.72	
G125	1.05	0.79	0.32
Y126	1.01	0.95	0.40
T127	0.96	0.97	0.54
N128	0.96	0.90	0.73
V129	0.98	0.85	0.63
P130			
V131	0.89	0.96	0.51
K132	0.90	1.00	
L133			
H134	0.98	1.00	0.67
P135			
E136	1.05	0.53	0.58
V137	1.05	0.93	0.94
T138	1.01	1.03	0.98
A139	0.92	1.04	0.90
T140	1.03	1.10	1.06
L141	0.83	0.37	0.95
K142	1.00		
V143	1.03	0.92	1.08
H144	0.93	1.02	0.96
V145	0.90	1.07	1.01
T146	0.92	1.08	1.04
E147	1.08	0.91	0.96
Q148	1.02	0.93	0.95
K149		1.08	1.07

Appendix 6. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of the pH 3.8 native state of CTL9

Residue	NH	N	CA	CB	CO	HA	HB	HB'
A58F			52.05	19.28	173.84	4.18	1.64	
A59F	8.75	124.13	53.61	19.28	178.24	4.42	1.50	
E60F	8.74	120.45	57.76	28.88	177.02	4.36	2.14	
E61F	8.51	120.12	58.86	29.30	178.63			
L62F			57.73	41.60	177.72	3.61	1.75	1.38
A63F	8.10	121.16	55.74	18.51	180.92	4.01	1.47	
N64F	8.29	116.66	56.15	38.13	177.70	4.49	2.90	
A65F	8.07	124.04	55.59	18.85	179.06	4.25	1.43	
K66F	8.30	117.88	60.83	32.54	179.20	3.85	1.89	
K67F	7.74	120.23	59.81	32.71	179.04	4.16	1.95	
L68F	7.99	121.77	57.77	42.15	178.55	4.28	1.75	
K69F	8.44	120.71	60.45	32.78	176.92	3.68	2.13	1.60
E70F	7.51	116.07	58.88	28.49	178.81	4.07	2.22	
Q71F	7.70	117.05	58.85	29.53	179.48	4.03	2.27	2.10
L72F	8.84	118.46	58.06	42.64	179.10	3.82	1.85	1.32
E73F	7.91	114.27	58.68	28.58	176.62	4.35	2.27	2.06
K74F	7.20	116.47	56.37	33.86	176.56	4.52	2.06	1.82
L75F	7.32	117.87	54.53	44.77	176.05	4.63	1.54	0.94
T76F	8.37	114.98	62.76	71.63	174.63	4.98	4.03	
V77F	8.94	127.12	61.74	34.04	174.06	4.63	2.09	
T78F	9.18	126.52	62.66	69.94	174.45	5.40	4.00	
I79F	8.94	128.42	58.67	42.31	172.31			
P80F			61.45	32.58	177.37	5.02	1.96	1.71
A81F	9.16	125.12	51.67	24.22	176.04	4.56	1.22	
K82F	8.40	121.45	57.07	33.59	175.67	4.52	1.82	
A83F	8.68	127.30	51.20	23.30	176.67	4.63	1.15	
G84F	8.18	108.23	44.01		174.46			
E85F	8.67	122.82	57.67	28.83	178.17	4.18	2.10	
G86F	9.00	112.04	46.32		175.15	4.00		
G87F	8.16	107.18	46.10		173.18			
R88F	7.12	118.82	56.30	30.60	176.54			
L89F			55.30	42.09	177.77	4.31	1.51	
F90F	7.91	120.21	59.20	39.12	176.63	4.56	3.23	2.84
G91F	8.28	109.53	44.74			4.28	3.89	
S92F	8.14	114.98	58.02	65.52				
I93F			60.56	38.58	175.26	4.52	1.88	
T94F	8.21	115.76	60.41	72.75	175.90	4.81	4.81	
S95F	9.32	116.38	58.20	63.14	175.36			
K96F			60.09	33.08	177.85	3.96	1.92	1.78
Q97F	7.51	117.95	60.09	29.27	179.50	4.24	2.42	2.19
I98F	7.96	120.12	66.48	38.32	176.94	3.57	1.89	
A99F	8.50	121.71	56.24	17.87	181.33	3.99	1.61	
E100F	8.52	119.06	59.89	28.89	179.54			
S101F	8.38	117.91	59.07	63.23	175.95	3.91	3.91	
L102F	8.59	122.94	59.07	42.30	178.75	4.07	2.13	1.64
Q103F	7.73	120.21	59.88	28.39	178.77	4.00	2.24	

A104F	8.38	122.20	55.71	19.18	179.95	4.16	1.62	
Q105F	8.46	112.01	57.52	29.89	177.77	4.21	1.92	1.64
H106F	7.87	111.17	54.85	31.52	174.49	5.26	3.58	3.02
G107F	8.11	109.91	47.28		173.13	4.07		
L108F	6.58	119.95	54.40	44.67	174.63	4.56	1.29	
K109F	8.64	125.82	56.16	33.42	174.25	4.33	1.82	1.64
L110F	8.68	128.10	53.30	45.96	175.16	4.56	1.50	1.22
D111F	8.19	123.88	54.79	42.40	177.61	4.45	2.69	
K112F	8.79	126.68	59.44	31.63	177.62			
R113F	8.67	119.65	57.79	30.07	177.87	4.17	1.88	
K114F	8.05	116.54	57.79	33.12	175.57	4.17	2.10	1.74
I115F	7.43	119.27	62.36	37.66	175.18	3.99	2.19	
E116F	8.63	129.72	54.95	28.08	175.29			
L117F	7.73	123.17	54.73	43.30	177.07	4.46	1.82	1.26
A118F	8.84	129.40	54.73	18.94	177.60	4.16	1.46	
D119F	7.77	113.93	52.72	41.51	173.94			
A120F			52.96	19.71	177.14	4.28	1.32	
I121F	8.54	121.71	63.21	39.29	176.93	4.00	1.78	
R122F	8.79	124.82	56.16	32.30	175.02	4.63	2.22	1.57
A123F	7.47	123.48	51.71	22.99	175.91	5.09	1.36	
L124F	8.10	117.66	55.58	43.50	177.13	3.89	1.75	1.47
G125F	8.58	107.30	44.75		174.50			
Y126F	9.14	122.84	58.93	40.79	176.36	5.53	2.87	
T127F	9.33	119.12	63.00	72.03	172.02	4.53	3.86	
N128F	8.77	125.42	52.20	39.61	174.65	5.39	2.79	2.38
V129F	9.40	125.82	59.57	34.70	172.73			
P130F			62.78	32.38	176.26	5.02	2.27	1.92
V131F	9.03	120.79	60.42	35.06	175.49	4.82	1.82	
K132F	9.11	129.61	56.02	32.06	175.49	4.62	1.92	1.64
L133F	8.06	125.96	56.35	43.85	176.41			
H134F								
P135F								
E136F								
V137F	8.18	119.80	63.07	35.14	173.49	4.49	1.96	
T138F	8.79	125.01	61.65	70.99	173.49	5.16	4.00	
A139F	9.39	128.41	49.65	24.13	173.41	4.95	1.35	
T140F	8.83	116.70	62.95	69.98	172.40	4.59	3.89	
L141F	9.12	128.09	53.22	45.84	175.29	4.69	1.92	
K142F	9.00	129.61	57.59	32.84	174.82	4.67	1.99	
V143F	8.97	126.44	61.47	34.13	174.54	4.84	2.13	
H144F	9.62	126.67	51.20	31.22	172.53	4.99	3.28	2.38
V145F	9.05	129.73	62.13	32.06				
T146F			59.32	70.83	172.25	4.38	3.79	
E147F	8.39	120.14	56.26	29.71	176.43	4.60	1.99	
Q148F	8.29	128.30	56.15	29.88	174.50	4.28	1.74	
K149F	8.36	131.37	57.75	33.65	175.18			

Appendix 7. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of the pH 3.8 unfolded state of CTL9

Residue	NH	N	CA	CB	CO	HA	HB	HB'
A58U			52.05	19.59	173.53	4.14	1.58	
A59U	8.64	123.70	53.18	19.59	177.71	4.37	1.44	
E60U	8.53	120.26	56.75	29.45	176.29	4.33	2.06	
E61U	8.46	121.65	56.68	29.40	176.39			
L62U	8.09	124.04	56.17	42.51	177.67	4.36	1.67	
A63U	8.29	124.03	53.69	19.16	178.24	4.24	1.43	
N64U	8.30	117.20	54.09	38.92	175.69	4.63	2.87	
A65U	8.14	123.77	54.45	19.01	178.72	4.21	1.49	
K66U	8.13	119.32	57.94	32.94	177.49	4.19	1.88	
K67U	8.03	120.89	57.76	32.97	177.53	4.23	1.86	
L68U	8.08	121.76	56.24	42.49	178.08	4.28	1.71	1.61
K69U	8.13	120.72	58.11	32.66	177.54	4.19	1.89	
E70U	8.18	179.71	57.73	29.24	177.03	4.25	2.11	
Q71U	8.25	120.27	57.33	29.27	176.83	4.30	2.12	
L72U	8.18	121.81	56.41	42.52	178.14	4.30	1.75	1.63
E73U	8.19	119.91	56.91	29.24	176.44	4.31	2.10	
K74U	8.08	120.98	57.18	32.92	176.58	4.30	2.22	
L75U	8.11	122.48	55.91	42.49	177.54	4.44	1.69	
T76U	8.11	115.25	62.45	70.06	174.35			
V77U	8.10	122.29	62.49	33.09	175.98	4.22	2.12	
T78U	8.22	118.24	62.49	70.08	174.05	4.38	4.14	
I79U	8.28	125.53	58.98	39.08	174.41			
P80U			63.47	32.46	176.63	4.42	2.29	1.93
A81U	8.33	124.25	53.09	19.69	177.91	4.30	1.40	
K82U	8.23	120.18	56.39	33.46	176.41	4.32	1.81	
A83U	8.31	125.15	53.39	19.62	178.24	4.32	1.43	
G84U	8.39	108.23	45.67		174.45	4.00		
E85U	8.25	119.86	56.39	29.45	176.87	4.40	2.18	2.03
G86U	8.55	109.88	46.09		174.82	3.99		
G87U	8.27	108.53	45.78		174.23	3.99		
R88U	8.10	120.29	56.56	31.36	176.24	4.32	1.79	
L89U	8.23	122.67	55.81	42.56	177.05	4.31	1.57	1.45
F90U	8.22	120.54	58.43	39.83	177.07	4.63	3.20	3.04
G91U	8.32	110.15	45.83		174.08	3.97		
S92U	8.18	115.63	58.96	64.48	174.68	4.53	3.92	
I93U	8.24	122.31	61.77	38.96	176.58	4.32	1.94	
T94U	8.18	116.97	62.25	70.17	175.00	4.45	4.32	
S95U	8.35	117.63	59.85	63.92	175.27	4.38	3.96	
K96U	8.26	123.15	57.65	33.07	177.09	4.27	1.85	
Q97U	8.18	120.32	57.66	29.49	177.17	4.27	2.10	
I98U	8.11	121.76	62.84	38.72	176.83	4.01	1.90	
A99U	8.25	125.77	54.11	19.05	178.85	4.24	1.45	
E100U	8.28	118.68	57.55	29.01	177.27	4.27	2.11	
S101U	8.20	116.25	59.90	63.94	175.36	4.42	3.99	
L102U	8.15	123.21	56.49	42.44	178.26	4.30	1.76	1.60
Q103U	8.11	119.40	57.32	29.45	176.57	4.24	2.09	
A104U	8.09	123.50	53.64	19.25	178.08	4.27	1.44	

Q105U	8.14	118.18	56.53	29.51	176.28	4.26	2.06	
H106U	8.39	118.27	56.10	29.40	174.97	4.73	3.38	3.26
G107U	8.40	109.38	45.95		174.01	4.00		
L108U	8.06	121.63	56.10	42.90	177.40	4.37	1.63	
K109U	8.35	122.18	56.43	33.04	176.50			
L110U	8.22	123.40	55.82	42.63	176.96	4.32	1.62	
D111U	8.36	120.84	54.23	40.42	176.08	4.64	2.81	
K112U	8.31	122.17	57.58	32.75	177.09	4.22	1.86	
R113U	8.24	120.57	57.76	30.92	177.05	4.24	1.88	
K114U	8.15	121.18	57.50	33.02	177.30	4.23	1.86	
I115U	8.00	121.77	62.42	38.98	176.54	4.07	1.90	
E116U	8.34	123.09	57.07	29.32	177.07	4.34	2.08	
L117U	8.24	123.06	56.60	42.58	177.72	4.30	1.69	
A118U	8.19	122.88	54.21	18.91	178.74	4.21	1.46	
D119U	8.27	117.58	55.04	39.52	176.31	4.57	2.85	
A120U	8.04	123.79	54.08	19.00	178.54	4.26	1.46	
I121U	7.97	118.47	62.97	38.45	177.29	4.00	1.88	
R122U	8.10	123.26	57.52	31.13	176.86			
A123U	8.30	122.89	53.27	19.18	178.05	4.29	1.44	
L124U	8.00	120.00	55.97	42.36	178.05	4.29	1.72	1.56
G125U	8.16	118.00	45.70		174.02	3.94		
Y126U	7.91	119.56	58.36	39.19	175.93	4.66	3.11	2.98
T127U	8.05	115.26	62.03	70.06	173.77	4.35	4.20	
N128U	8.36	121.17	53.45	39.44	174.56	4.74	2.79	
V129U	8.04	121.77	60.24	32.96	174.31			
P130U								
V131U	8.16	120.76	63.14	33.43	176.11	4.06	2.05	
K132U	8.32	125.05	56.30	33.43	176.11			
L133U	8.20	123.24	56.60	42.66	176.74	4.30	1.61	1.44
H134U	8.47	118.77	53.44	28.81	172.27			
P135U			63.58	32.45	176.83	4.44	2.31	1.96
E136U	8.66	121.09	56.25	29.56	176.32			
V137U	8.26	121.86	62.43	33.17	176.11	4.21	2.09	
T138U	8.32	118.68	62.43	70.05	174.23	4.35	4.21	
A139U	8.38	126.83	53.21	19.80	177.79	4.40	1.43	
T140U	8.11	113.69	62.41	70.03	174.46	4.32	4.18	
L141U	8.18	124.80	55.76	42.65	177.07	4.39	1.62	
K142U	8.30	122.73	56.34	33.20	176.32	4.37	1.73	
V143U	8.02	120.88	62.57	33.17	175.78	4.10	2.03	
H144U	8.65	122.69	55.49	29.38	174.17	4.79	3.23	
V145U	8.30	122.73	62.93	33.28	176.31	4.26	2.12	
T146U	8.28	119.01	62.37	70.08	174.05	4.37	4.20	
E147U	8.43	123.39	56.31	29.76	175.69	4.43	2.13	1.99
Q148U	8.45	122.61	56.21	29.83	175.18	4.36	2.14	2.00
K149U	8.20	127.21	57.49	33.70	180.35			

Appendix 8. ^{15}N R2 relaxation rates for the pH 3.8 native state and the pH 3.8 unfolded states of CTL9 and calculated R2 rates using random coil model

Residue	R_2 pH 3.8 native (s^1)	R2 pH 3.8 unfolded (s^1)	Model value (s-1)
A59	5.27±0.37	2.37±0.11	1.76
E60	4.38±0.14	1.39±0.04	1.91
E61	7.59±0.27	2.46±0.05	2.04
L62			2.15
A63	5.89±0.14	3.18±0.11	2.25
N64	7.18±0.16	2.09±0.05	2.34
A65		3.36±0.07	2.42
K66		3.48±0.06	2.49
K67		3.12±0.09	2.55
L68			2.60
K69		3.54±0.07	2.65
E70	10.59±0.33	3.72±0.08	2.69
Q71	10.72±0.40		2.73
L72	10.24±0.29	4.23±0.10	2.76
E73	10.95±0.33		2.79
K74	10.28±0.43		2.81
L75	10.57±0.36	3.38±0.08	2.84
T76		4.03±0.06	2.86
V77	14.30±0.45		2.87
T78	11.54±0.31	3.19±0.10	2.89
I79	12.45±0.48		2.90
P80			2.91
A81	9.62±0.35	2.34±0.13	2.92
K82			2.93
A83	9.98±0.31	4.96±0.13	2.94
G84		3.19±0.05	2.95
E85		2.78±0.05	2.96
G86	8.00±0.24	3.71±0.07	2.96
G87	9.11±0.22	2.60±0.07	2.97
R88	12.64±0.48	3.19±0.07	2.97
L89		2.74±0.09	2.97
F90	6.74±0.24	2.86±0.15	2.98
G91	7.09±0.14		2.98
S92	6.10±0.18	2.46±0.05	2.98
I93		2.81±0.09	2.98
T94	8.54±0.47	4.14±0.06	2.99
S95	11.17±0.47	4.04±0.07	2.99
K96		3.11±0.47	2.99
Q97	11.23±0.44	4.56±0.09	2.99
I98	8.26±0.37	4.04±0.05	2.99
A99		5.20±0.15	2.99
E100	8.81±0.20	5.36±0.07	2.99
S101		4.42±0.05	2.99
L102	8.84±0.21	4.24±0.08	2.99
Q103	8.48±0.24		2.99
A104		4.49±0.10	2.99
Q105	7.69±0.26	3.80±0.06	2.99

H106	8.94±0.42	3.58±0.04	2.99
G107	8.74±0.35	4.28±0.08	2.99
L108	10.30±0.60	2.47±0.07	2.99
K109	10.57±0.30	3.96±0.07	2.99
L110	9.93±0.48		2.99
D111	5.66±0.28		2.99
K112	12.00±0.50	4.63±0.13	2.99
R113		3.17±0.04	2.99
K114	9.00±0.24	5.95±0.11	2.98
I115	9.60±0.60	5.44±0.11	2.98
E116	8.22±0.30	5.45±0.17	2.98
L117			2.98
A118		4.98±0.11	2.97
D119	4.00±0.70	4.55±0.06	2.97
A120		4.45±0.11	2.97
I121		5.14±0.07	2.96
R122	13.10±0.50		2.96
A123			2.95
L124	9.84±0.37	4.96±0.15	2.94
G125	8.56±0.35	4.97±0.10	2.93
Y126		3.94±0.04	2.92
T127	9.34±0.28	3.95±0.07	2.91
N128	12.20±0.70	3.10±0.07	2.90
V129	12.90±0.60	2.72±0.06	2.89
P130			2.87
V131	10.20±0.48		2.86
K132	9.64±0.47	3.93±0.16	2.84
L133	7.17±0.19	3.25±0.39	2.81
H134		3.28±0.04	2.79
P135			2.76
E136		3.15±0.13	2.73
V137		3.15±0.05	2.69
T138	4.98±0.33	1.79±0.05	2.65
A139	10.02±0.30	3.63±0.04	2.60
T140	9.90±0.32	1.88±0.05	2.55
L141	9.30±0.44	4.28±0.08	2.49
K142	10.31±0.41		2.42
V143		3.12±0.09	2.34
H144	10.70±0.60	2.01±0.07	2.25
V145	9.75±0.47		2.15
T146		3.92±0.05	2.04
E147	7.02±0.21	1.33±0.06	1.91
Q148	8.63±0.27	1.83±0.05	1.76
K149	6.20±0.25	2.59±0.06	1.60

Appendix 9. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of I98A CTL9 at pH 5.7 and 12 ^{o}C

Residue	NH	15N	CA	CB	CO	HA	HB	HB'
A58			52.26	19.15	173.64	4.13	1.51	--
A59	8.67	123.58	53.41	19.15	177.97	4.32	1.37	--
E60	8.69	120.71	57.58	30.12	176.81	4.23	1.93	--
E61	8.48	121.82	57.50	30.15	177.20	4.26	1.94	--
L62	8.19	123.05	56.12	42.33	178.04	4.34	1.59	--
A63	8.24	123.80	53.99	18.94	178.71	4.20	1.37	--
N64	8.30	117.29	54.29	38.70	175.93	4.62	2.78	--
A65	8.07	123.62	54.47	18.90	179.03	4.20	1.41	--
K66	8.08	119.61	58.26	32.85	177.60	4.19	1.80	--
K67	7.97	120.93	57.76	32.83	177.76	4.33	1.73	--
L68	8.06	121.61	56.15	42.56	178.36	4.26	1.60	--
K69	8.04	120.75	58.30	32.88	177.75	4.16	1.84	--
E70	8.18	120.04	58.20	30.10	177.69	4.19	1.99	--
Q71	8.15	119.87	57.27	29.32	177.13	4.25	2.08	--
L72	8.12	121.77	56.68	42.41	178.55	4.27	1.62	--
E73	8.19	120.39	57.60	30.15	176.99	4.20	1.98	--
K74	7.99	120.70	57.14	32.83	176.80	4.27	1.79	--
L75	8.03	122.05	55.73	42.59	177.64	4.41	1.58	--
T76	8.08	115.52	62.26	69.90	174.31	4.27	4.19	--
V77	8.12	123.08	62.25	33.15	176.04	4.20	2.01	--
T78	8.31	119.70	62.11	69.97	174.06	4.34	4.05	--
I79	8.33	126.15	58.71	38.80	174.31	--	--	--
P80			63.26	32.35	176.60	4.41	2.22	1.83
A81	8.37	124.58	52.66	19.45	177.82	4.27	1.30	--
K82	8.32	120.92	56.21	33.39	176.37	4.27	1.72	--
A83	8.38	125.72	53.18	19.28	178.31	4.27	1.30	--
G84	8.41	108.59	45.51	--	174.37	3.99	--	--
E85	8.25	120.64	57.11	30.35	177.27	4.34	1.97	--
G86	8.57	110.12	46.03	--	174.84	3.98	--	--
G87	8.27	108.60	45.68	--	174.25	3.96	--	--
R88	8.05	120.36	56.45	31.04	176.24	4.27	1.68	--
L89	8.25	122.90	55.32	42.47	177.15	4.33	1.42	--
F90	8.26	120.99	58.26	39.87	176.30	4.62	3.02	--
G91	8.33	110.49	45.60	--	174.02	3.91	--	--
S92	8.14	115.64	58.38	64.21	174.64	4.53	3.82	--
I93	8.28	122.47	61.81	38.90	176.70	4.27	1.86	--
T94	8.21	117.43	62.03	69.96	174.91	4.34	4.26	--
S95	8.35	118.28	59.34	63.71	175.19	4.41	3.84	--
K96	8.32	123.59	57.62	32.78	177.18	4.24	1.79	--
Q97	8.23	120.95	57.09	29.32	177.05	4.27	2.00	--
A98	8.32	124.85	53.58	19.02	178.47	--	--	--
A99			53.69	19.00	178.85	4.24	1.38	--
E100	8.33	119.38	57.81	30.14	177.57	4.20	2.01	--
S101	8.17	116.19	59.50	63.61	175.32	4.41	3.84	--
L102	8.09	123.46	56.36	42.35	178.20	4.27	1.58	--
Q103	8.12	119.59	56.87	29.18	176.54	4.23	2.00	--
A104	8.09	123.79	53.43	19.08	178.20	4.27	1.36	--

Q105	8.15	118.38	56.44	29.32	176.30	4.27	1.94	--
H106	8.34	118.69	56.09	29.40	175.19	4.69	3.20	--
G107	8.38	109.45	45.70	--	174.00	3.99	--	--
L108			56.36	42.75	177.40	4.34	1.53	--
K109	8.34	122.40	56.36	32.87	176.48	4.33	1.72	--
L110	8.23	123.76	55.74	42.58	176.79	4.33	1.52	--
D111	8.33	121.69	54.70	41.40	176.48	4.60	2.63	--
K112	8.30	122.49	57.25	32.80	176.82	--	--	--
R113	8.23	120.95	57.09	30.76	177.05	4.24	1.79	--
K114	8.15	121.68	57.27	32.96	177.10	4.27	1.79	--
I115	8.03	122.54	62.04	38.86	176.55	4.07	1.80	--
E116	8.44	124.17	57.16	30.25	177.48	--	--	--
L117			56.25		178.49	4.25	1.37	--
A118	8.19	122.68	53.66	19.01	178.83	4.20	1.38	--
D119	8.21	118.92	55.82	40.99	177.15	4.50	2.63	--
A120	8.03	123.75	53.92	18.98	178.91	4.20	1.38	--
I121	7.94	118.58	62.76	38.33	177.63	3.95	1.84	--
R122	8.05	123.29	57.49	30.72	177.17	4.20	1.79	--
A123	8.05	123.04	53.39	19.03	178.26	4.20	1.37	--
L124	7.92	119.94	55.94	42.43	178.22	4.27	1.65	1.45
G125	8.10	107.91	45.63	--	174.02	3.91	--	--
Y126	7.87	119.69	58.23	38.91	175.93	4.62	2.99	--
T127	8.02	115.66	61.88	70.05	173.60	4.34	4.11	--
N128	8.35	121.35	53.45	39.10	174.55	4.69	2.69	--
V129	8.05	122.29	60.45	32.80	174.36	--	--	--
P130			62.74	32.87	176.78	4.41	2.22	1.78
V131	8.21	121.22	62.74	32.87	176.13	4.05	1.94	--
K132	8.36	125.67	56.18	33.37	176.09	4.34	1.65	--
L133	8.28	123.85	55.52	42.73	176.72	4.18	1.51	1.31
H134	8.46	118.74	53.54	29.00	172.43	--	--	--
P135			63.42	32.33	176.84	4.41	2.23	1.88
E136	8.86	121.54	57.04	30.10	176.69	4.34	1.93	
V137	8.24	122.15	62.76	32.83	176.47	4.18	2.07	
T138	8.19	118.51	62.24	69.82	174.34	4.27	4.19	
A139	8.35	126.83	53.11	19.48	177.80	4.40	1.37	
T140	8.11	114.04	62.16	69.81	174.43	4.26	4.12	
L141	8.18	125.07	55.51	42.55	177.06	4.33	1.52	
K142	8.31	123.15	56.10	33.21	176.17	4.33	1.72	
V143	8.07	121.60	62.21	33.13	175.68	4.09	1.93	
H144	8.64	123.38	55.65	29.62	174.33	4.74	3.13	
V145	8.30	123.38	62.40	33.23	176.12	4.21	2.00	
T146	8.31	118.94	62.22	69.83	174.24	4.30	4.12	
E147	8.49	124.26	56.75	30.64	175.93	4.34	1.94	
Q148	8.46	122.96	56.13	29.86	174.88	4.34	1.97	
K149	8.10	129.21	57.86	33.86	181.29			

Appendix 10. Backbone and $^{13}C_{\beta}$ and $^{1}H_{\beta}$ assignments of wild type CTL9 at pH 5.7 and 12 ^{o}C

Residue	NH	N15	CA	CB	CO	HA
A59	9.06	124.19	54.01	18.84	178.92	4.22
E60	9.20	120.42	58.74	29.29	177.69	3.73
E61	8.70	119.53	59.77	29.41		
L62						
A63			55.81	18.13	181.61	3.73
N64	8.66	116.98	55.81	37.87	178.15	4.43
A65	8.22	124.29	55.36	18.48	179.33	4.23
K66	8.53	118.18	60.84	32.45		
K67			59.88	29.46	179.73	4.03
L68	8.21	122.44	58.15	41.57	178.62	4.24
K69	8.60	121.06	60.29	32.62	176.78	3.33
E70	7.58	116.23	59.09	29.55	179.28	3.85
Q71	7.77	116.58	58.94	29.39	179.74	3.81
L72	9.14	118.50	57.72	42.78	179.54	3.49
E73	8.17	114.41	58.83	28.49		
K74	7.31	116.65	56.49	33.95	176.84	4.50
L75	7.45	118.10	54.23	44.28	176.36	4.63
T76	8.50	115.19	62.50	71.55	174.84	5.03
V77	9.20	128.87	61.46	33.40	174.32	4.61
T78	9.59	127.42	62.39	69.32		
I79						
P80			61.01	32.36	177.81	4.34
A81	9.41	125.27	51.49	24.31	176.24	4.53
K82	8.66	122.13	56.86	33.23	176.07	4.50
A83	9.30	128.07	50.94	23.50	176.53	4.68
G84	8.36	108.35	43.71		174.39	4.51
E85	8.90	123.51	57.92	29.76	178.91	3.99
G86	9.30	112.65	46.02		175.39	3.70
G87	8.37	107.33	45.86		173.31	
R88	7.18	118.69	55.81	30.28	176.73	4.33
L89	8.26	124.20	55.31	41.57	178.96	4.24
F90	7.98	120.19	59.42	38.93	176.94	4.50
G91	8.57	109.60	44.25		172.07	3.53
S92					173.38	
I93					175.38	4.51
T94	8.49	116.03	59.94	72.42	176.03	4.91
S95	9.59	116.46	59.93	72.39	176.00	
K96						
Q97			59.58	29.08	179.70	4.17
I98	8.18	120.08	66.28	37.98	177.05	3.16
A99	8.68	121.70	55.95	17.81	177.88	3.73
E100	7.61	117.60	59.73	29.08	180.26	4.07
S101	8.64	118.59	59.66	63.04	175.74	4.43
L102	8.85	123.28	59.04	42.14	178.72	3.92
Q103	7.88	119.55	59.34	28.13	178.98	3.75
A104	8.38	121.97	55.35	19.09	180.05	3.99
Q105	8.68	113.37	57.95	29.41	177.74	

H106	7.91	112.89	55.67	33.70	176.32	5.37
G107	8.18	109.62	46.76		173.46	
L108	6.68	119.87	53.94	44.29	174.82	4.53
K109	8.90	125.85	55.82	33.09	174.37	4.19
L110	8.93	128.81	52.89	45.89	175.18	4.51
D111	8.40	124.16	54.36	42.37	177.79	4.34
K112	9.06	126.99	59.10	31.50	177.77	3.65
R113	8.91	120.06	57.98	29.97	178.29	3.99
K114	8.26	116.60	57.89	33.06	175.59	4.06
I115	7.63	119.28	61.99	37.58	179.70	3.55
E116	7.93	120.41	59.21	32.49	179.17	
L117			58.15	41.58	175.13	3.82
A118	8.22	121.01	55.43	18.11		
D119					177.83	5.02
A120					181.56	
I121			63.30	38.91	177.23	3.73
R122	9.02	124.64	55.66	32.13	175.28	4.68
A123	7.63	123.64	51.31	23.16	176.01	5.35
L124	8.27	117.79	55.24	43.41	177.40	3.64
G125	8.84	107.52	44.37		174.95	4.68
Y126	9.44	123.17	58.63	40.23	176.69	
T127	9.58	119.53	62.39	71.83	172.19	4.50
N128	9.00	125.27	52.08	39.59		
V129						
P130			62.34	34.43	176.51	5.19
V131	9.35	121.23	60.29	34.43	173.72	4.87
K132	9.44	130.11	55.56	31.24	175.80	4.67
L133	8.21	126.21	55.47	44.03	176.48	4.14
H134	8.76	121.53	55.48	34.52	173.19	
P135						
E136			55.72	31.28	176.94	4.63
V137	8.35	119.77	62.57	35.54		
T138			61.47	70.60	173.47	4.50
A139	9.66	128.93	49.31	23.80	175.59	5.11
T140	9.68	118.35	62.49	69.43	172.72	5.36
L141	9.41	129.30	52.87	45.52	174.76	4.68
K142	9.30	130.05	57.50	32.74	174.95	4.68
V143	9.12	126.38	60.89	33.98	174.64	4.96
H144	9.82	127.19	51.14	31.58	172.79	5.11
V145			61.86	31.73	176.23	5.02
T146	8.25	121.06	59.06	70.34	172.45	4.16
E147	8.66	120.25	56.73	30.62	176.75	4.37
Q148	8.39	128.73	55.22	30.20	174.47	4.20
K149	8.56	133.29	57.68	33.47	181.53	

Appendix 11. ^{15}N R_2 relaxation rates for I98A CTL9, wt-CTL9 at pH 5.7, 12 ^{o}C and calculated R_2 rates using random coil model

Residue	R_2 I98A CTL9 (s^1)	R_2 wt-CTL9 (s^1)	Model value (s-1)
A59	4.41±0.14		1.76
E60	2.59±0.14	8.95±0.10	1.91
E61	3.46±0.09	14.57±0.09	2.04
L62	5.67±0.10		2.15
A63	5.04±0.11		2.25
N64	5.53±0.12	15.60±0.15	2.34
A65	5.91±0.16	16.21±0.34	2.42
K66	5.56±0.16	18.29±0.15	2.49
K67	7.32±0.06		2.55
L68	4.79±0.10	16.32±0.26	2.60
K69	7.00±0.11	17.43±0.25	2.65
E70	6.62±0.14	17.73±0.21	2.69
Q71	7.42±0.13	16.54±0.32	2.73
L72	7.43±0.15	17.81±0.28	2.76
E73	7.04±0.15	17.56±0.22	2.79
K74	7.22±0.16	15.87±0.16	2.81
L75	6.21±0.16	15.59±0.19	2.84
T76	5.78±0.12	14.48±0.15	2.86
V77	5.18±0.11	16.21±0.39	2.87
T78	5.91±0.13	17.01±0.41	2.89
I79	5.47±0.18		2.90
P80			2.91
A81	3.58±0.25	13.39±0.17	2.92
K82	3.70±0.11	13.37±0.11	2.93
A83	4.67±0.10	15.28±0.31	2.94
G84	4.65±0.09	15.98±0.20	2.95
E85	4.10±0.12	12.38±0.12	2.96
G86	4.65±0.10	14.23±0.19	2.96
G87	4.28±0.12	16.17±0.34	2.97
R88	4.53±0.11	17.73±0.20	2.97
L89	3.92±0.11	15.41±0.40	2.97
F90	3.76±0.16	13.54±0.26	2.98
G91	4.31±0.12	10.10±0.24	2.98
S92	4.92±0.10	9.37±0.09	2.98
I93	5.05±0.16		2.98
T94	5.89±0.14	18.40±1.00	2.99
S95	6.32±0.10	17.50±0.90	2.99
K96			2.99
Q97	7.70±0.16		2.99
A98/I98	6.02±0.25	15.18±0.15	2.99
A99		16.18±0.21	2.99
E100	5.64±0.13	21.40±1.40	2.99
S101	6.56±0.11	16.11±0.16	2.99
L102	5.35±0.17	14.89±0.16	2.99
Q103	6.56±0.13	15.66±0.17	2.99
A104	5.90±0.18	15.03±0.15	2.99
Q105	5.53±0.15	14.30±0.16	2.99

H106	7.06±0.14	13.59±0.26	2.99
G107	7.41±0.09	19.76±0.35	2.99
L108		15.68±0.38	2.99
K109	5.63±0.09	14.04±0.14	2.99
L110		17.67±0.24	2.99
D111	5.67±0.19	15.97±0.25	2.99
K112		18.76±0.36	2.99
R113		15.86±0.20	2.99
K114	10.59±0.30	15.11±0.28	2.98
I115	7.18±0.24	16.69±0.42	2.98
E116	10.68±0.26	15.90±0.10	2.98
L117			2.98
A118	5.95±0.12	15.40±1.40	2.97
D119	7.71±0.23		2.97
A120	7.13±0.25		2.97
I121	8.37±0.16		2.96
R122	8.26±0.17	17.90±0.70	2.96
A123	8.10±0.16	18.64±0.47	2.95
L124	8.69±0.13	18.18±0.41	2.94
G125	7.74±0.19	15.98±0.17	2.93
Y126	6.31±0.14	15.65±0.28	2.92
T127	6.15±0.12	15.10±0.80	2.91
N128	5.55±0.14	21.60±2.10	2.90
V129	4.71±0.10		2.89
P130			2.87
V131	4.64±0.13	15.80±0.70	2.86
K132	5.20±0.13	15.89±0.23	2.84
L133	4.00±0.15	15.10±1.30	2.81
H134	5.76±0.10	15.88±0.19	2.79
P135			2.76
E136	4.39±0.10		2.73
V137	3.76±0.16	13.83±0.28	2.69
T138	4.68±0.12		2.65
A139	6.97±0.23	15.70±0.60	2.60
T140	4.84±0.11	15.63±0.29	2.55
L141	4.01±0.29	14.60±0.60	2.49
K142	4.36±0.11	16.90±0.50	2.42
V143		15.82±0.50	2.34
H144	5.36±0.09	19.20±0.50	2.25
V145	3.72±0.09	16.50±1.30	2.15
T146	3.74±0.12	16.06±0.28	2.04
E147	3.19±0.16	14.66±0.16	1.91
Q148	2.08±0.07	14.01±0.24	1.76
K149	2.31±0.13	9.45±0.14	1.60

Appendix 12. ^{1}H-^{15}N RDCs for I98A CTL9 and wt-CTL9 at pH 5.7, 12 ^{o}C

Residue	R_{NH} I98A CTL9	R_{NH} wt-CTL9
A59	-8.92	-14.09
E60	3.41	3.66
E61	2.95	5.12
L62	0.59	--
A63	--	--
N64	4.67	5.18
A65	4.87	12.87
K66	5.62	20.93
K67	7.50	--
L68	3.54	34.93
K69	3.63	26.47
E70	5.94	23.28
Q71	5.41	11.74
L72	3.61	11.61
E73	5.34	28.02
K74	0.73	5.00
L75	0.33	4.93
T76	-2.08	-6.87
V77	-3.69	-12.94
T78	-4.04	-7.34
I79	-6.63	--
P80	--	--
A81	-5.04	-9.56
K82	-4.46	-14.96
A83	-1.81	-4.51
G84	-0.61	0.38
E85	1.53	-3.58
G86	1.56	4.22
G87	1.58	11.24
R88	-0.75	20.66
L89	1.00	6.38
F90	2.56	-8.46
G91	2.47	-7.59
S92	-2.08	-8.33
I93	0.23	--
T94	0.97	
S95	0.89	-13.15
K96	8.92	--
Q97	--	--
98	2.51	-10.77
A99	--	-13.12
E100	0.78	-0.01
S101	0.94	-6.74
L102	1.67	-9.65
Q103	3.11	-7.57
A104	2.43	-8.25
Q105	1.20	-6.82
H106	-0.48	-0.21

G107	0.33	17.51
L108	--	16.28
K109	1.00	11.07
L110	2.10	16.13
D111	1.33	16.25
K112	-0.52	-15.28
R113	2.72	-0.80
K114	1.49	-6.42
I115	0.50	-12.02
E116	3.96	7.91
L117	--	--
A118	1.93	-9.50
D119	14.08	--
A120	8.12	--
I121	6.10	--
R122	4.64	-11.08
A123	8.58	16.14
L124	7.02	16.75
G125	-1.14	-12.22
Y126	7.85	-3.03
T127	-1.40	-12.98
N128	-1.07	-2.96
V129	-0.92	--
P130	--	--
V131	-1.46	-12.38
K132	-4.51	-5.43
L133	-1.15	-12.25
H134	-3.34	-0.89
P135	--	--
E136	-3.69	--
V137	-2.73	-13.54
T138	-3.95	--
A139	0.14	-11.47
T140	-2.49	-4.06
L141	-1.25	-9.66
K142	-2.22	-0.06
V143	1.36	-9.08
H144	-5.13	-10.96
V145	-4.56	-10.04
T146	-4.17	-11.00
E147	-3.32	-16.64
Q148	-3.22	-10.86
K149	-2.93	4.02