Stony Brook University

The Other Half of the Protein Folding Equation

Characterization of the Protein Denatured State Ensemble (DSE):

The Case of C-terminal Domain of the Ribosomal Protein L9 (CTL9)

A Dissertation Presented

by

Bowu Luan

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Chemistry

Stony Brook University

December 2014

Stony Brook University

The Graduate School

Bowu Luan

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., Dissertation Advisor
Professor, Department of Chemistry, Stony Brook University

David F. Green, Ph.D., Chairperson
Professor, Department of Chemistry, Department of Applied mathematics and Statistics, Stony Brook University

Kathlyn A. Parker, Ph.D., Third Member
Professor, Department of Chemistry, Stony Brook University

Mark E. Bowen, Ph.D., Outside Member
Professor, Department of Physiology and Biophysics, Stony Brook University

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

The Other Half of the Protein Folding Equation

Characterization of the Protein Denatured State Ensemble (DSE):

The Case of C-terminal Domain of the Ribosomal Protein L9 (CTL9)

by

Bowu Luan

Doctor of Philosophy

in

Chemistry

Stony Brook University

2014

Studies on the unfolded state of proteins are important for understanding the mechanism of protein folding. The protein unfolded state or protein denatured state ensemble (DSE) is the starting point for protein folding. It provides a thermodynamic reference state for protein stability, and it can be targeted by rational protein design. Studies of the properties of DSEs can also reveal the factors that impact protein misfolding and modulate the tendency for protein aggregation *in vitro* and *in vivo*, and amyloid formation.

The 92-residue C-terminal domain of ribosomal protein L9 (CTL9) is used as a model to study protein DSEs. This globular protein follows a two-state folding and unfolding transition. It includes three His residues which cause it to unfold at low pH. In addition to acid-induced denaturation, CTL9 can also be unfolded by adding chemical denaturants, such as urea. A core

mutant I98A-CTL9, that destabilizes the protein without perturbing the fold, is used to study protein cold denaturation and heat denaturation. The nature of temperature dependent unfolding remains controversial.

The aims of the research described in my thesis are to characterize the structural and dynamic properties of CTL9 DSEs populated under various denaturing conditions, study the folding cooperativity of cold denaturation, and compare the properties of the temperature dependent I98A-CTL9 DSE to the cold unfolded DSE. Nuclear Magnetic Resonance (NMR) was used to provide residue-specific information. Small angle X-ray scattering (SAXS) was used to monitor the overall shape and compactness. The acid- and urea-induced DSE of CTL9 showed significantly different long-range contacts within the unfolded ensemble, despite the fact that they have similar radius of gyration (R_g). The properties of the DSE populated under conditions where the folded state is also present were examined as a function of pH and urea, respectively. The two-state, cooperative unfolding of cold denaturation has been examined and confirmed using the core destabilizing point mutant I98A-CTL9. The properties of the cold- and heat-induced I98A-CTL9 DSEs are compared in terms of compactness and the tendencies to form secondary structure.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CD	circular dichroism
C_M	midpoint concentration of denaturant induced unfolding transition
CTL9	the C-terminal domain of protein L9
DSE	denatured state ensemble
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
ER	endoplasmic reticulum
EV	excluded volume
F	the folded state
HPLC	high performance liquid chromatography
HSQC	heteronuclear single coherence
I	the intermediate state
IDPs	intrinsically disordered proteins
IDRs	intrinsically disordered regions
LB	Luria-Bertani
M	molarity
MALDI	matrix assisted laser desorption and ionization
mM	millimolarity
MD	molecular dynamics
N	the native state
NTL9	the N-terminal domain of protein L9
NMR	nuclear magnetic resonance

PDB	protein data bank
ppm	parts per million
PRE	paramagnetic relaxation enhancement
R	universal gas constant
R_g	radius of gyration
R_h	radius of hydration
SASA	solvent assessable surface area
SAXS	small angle X-ray scattering
smFRET	single molecule Forster resonance energy transfer
SSP	secondary structure propensity
T	temperature
TFA	trifluoroacetic acid
T_m	midpoint of thermal unfolding transition
T_s	temperature of maximal stability
U	the unfolded state
UPR	unfolded protein response
UV	ultraviolet
WT	wild type
ΔC_p^o	heat capacity change upon unfolding
ΔG^o_{N-D}	standard Gibbs free energy of unfolding
ΔH^o	standard enthalpy of unfolding
ΔS^o	standard entropy of unfolding
[den]	concentration of denaturant

$[\theta]$	mean residue ellipticity
3D	three dimensional

Acknowledgments

I would like to thank my advisor Prof. Daniel Raleigh for guiding me into the field of protein folding and his support and advices during my thesis research. I am so grateful for his encouragement and patience. In addition, I learned a lot from him on scientific writing and presentation, together with many other transferable skills that are beneficial to me in developing my future career. I am lucky to have him as my advisor.

I would like to thank Professor David Green, Professor Kathlyn Parker, and Professor Mark Bowen for serving as my dissertation committee members. I thank them for their kind suggestions and advices.

I am grateful for the friendship with the members of Raleigh group: Dr. Bing Shan, Dr. Humeyra Taskent, Dr. Fanling Meng, Dr. Wenli Meng, Dr. Shifeng Xiao, Dr. Peter Marek, Dr. Vadim Patsalo, Dr. Ling-Hsien Tu, Dr. Ping Cao, Dr. Hui Wang, Ivan Peran, Natalie Stenzoski, Matthew Watson, Harris Norr, Rehana Akter, Xiaoxue Zhang, Amy Wong, Junjie Zou, Zack Ridgeway. I would like to thank Dr. Bing Shan, Dr. Wenli Meng, and Dr. Shifeng Xiao for introducing me to various experimental techniques and Ivan Peran and Natalie Stenzoski for productive collaborations.

I would like to thank my friends who have been helping me a lot during the years of my PhD study.

Last but not the least, I would like to thank my parents for their unconditional love, visionary insights, and everlasting support.

List of Publications

Lahiri, T., **Luan, B.**, Raleigh, D. P., and Boon, E. M. (2014) A structural basis for the regulation of an H-NOX-associated cyclic-di-GMP synthase/phosphodiesterase enzyme by nitric oxide-bound H-NOX. *Biochemistry, 53*: 2126-2135.

Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2014) Denatured state ensembles with the same radius of gyration can form significantly different long-range contacts. *Biochemisty, 53*: 39-47.

Luan, B., Shan, B., Baiz, C., Tokmakoff, A, and Raleigh D. P. (2013) Cooperative cold denaturation: The case of the C-terminal domain of the ribosomal protein L9. *Biochemistry, 52*: 2402-2409.

Meng, W., Lyle, N., **Luan, B.**, Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. *Proc. Natl. Acad. Sci. U.S.A., 110*: 2123-2128.

Meng, W., **Luan, B.**, Lyle, N., Pappu, R. V., and Raleigh, D. P. (2013) The denatured state ensemble contains significant local and long- range structure under native conditions: Analysis of the N-Terminal domain of ribosomal protein L9. *Biochemistry, 52*: 2662-2671.

Patsalo, V., Yondaola, M. A., **Luan, B.**, Shoshani, I., Kisker, C., Green, D. F., Raleigh, D. P., and Hearing, P. (2012) Biophysical and functional analyses suggest that adenovirus E4-ORF3 protein requires higher-order multimerization to function against promyelocytic leukemia protein nuclear bodies. *J. Bio. Chem., 287*: 22573-22583.

Chapter 1

Introduction

1.1. Protein folding

1.1.1. The importance of protein folding

Proteins are among the most important molecules in cells and organs, and have a wide range of functions. Protein folding is the process by which a newly synthesized unstructured polypeptide folds into a stable, functional and specific three dimensional structure after translated from mRNA on the ribosome. Protein folding is a critical step for transferring the genetic information from the DNA to the protein, and it is one of the most challenging and important problems in current chemical and molecular biology. Protein structure is important for function. Proteins that fail to fold correctly (misfolded protiens) can induce protein aggregation. The accumulation of the resulting deposits formed by the misfolded proteins is believed to be a leading contributor to neurodegenerative and neuromuscular diseases such as Parkinson's disease, Alzheimer's disease, Huntington's disease and can contribute to the disorders such as type-2 diabetes.[1-3] In addition, the accumulation of the unfolded and misfolded proteins in the endoplasmic reticulum (ER) can trigger the unfolded protein response (UPR), which plays a critical role in tumorigenesis.[4] A better understanding of the principles of protein folding will help in the explanation of how such diseases arise and assist in the development of therapeutic insights for these diseases caused by protein misfolding and

aggregation. A better understanding of protein folding will also aid efforts to design proteins with important properties and efforts to optimize protein production in biotechnology.

1.1.2. The mechanism of protein folding

Over the past half-century, the exploration of the major protein folding problem started as four questions: 1) How does amino acid sequence dictate the protein's native structure? 2) What determines protein folding rates and how to describe folding kinetics? [5] 3) Is it possible to design a model that is able to predict protein structure based on the sequence? [6-8] 4) Is protein folding different *in vivo* compared with that *in vitro*? [9, 10]

As demonstrated during the 1950s, the well-defined folded protein structure, also termed the protein native state or the folded state, is dictated by the amino acids in the polypeptide sequence. In addition, it was shown that unfolded proteins can refold spontaneously, given suitable conditions. This principle still holds even after the discovery of molecular chaperones and has inspired many studies in the field of protein folding ever since.[11] No further details were provided about the mechanism of how proteins find their correct fold from these important early studies.[12, 13]

In 1968, Prof. Cyrus Levinthal pointed out that despite the enormous possible conformations that the polypeptide can adopt, proteins fold precisely into their well-defined native structure fast, from the millisecond to the second timescale, with some

ultrafast folders in the microsecond scale.[14-16] Levinthal pointed out that a random search through all possible conformations would take an extraordinary long time. The question arises of how proteins manage to reduce the search time to achieve their native folded structures over numerous other possible conformations. Levinthal suggested that this paradox can be solved by the existence of preferred folding pathways through which the polypeptide can be guided to find its correct fold in such a high efficiency.[17]

Many models have been proposed to describe protein folding pathways, and a recent model based on statistical mechanics, suggests that the protein folding energy landscape is funnel-shaped.[18-21] The shape of the folding funnel is unique to different proteins and is determined by the amino acid sequence. There are fewer conformations of the more compact, low-energy ensembles in the funnel bottom, while on top of the folding funnel, there are many more unfolded structures with a high degree of configuration entropy. Instead of having a smooth surface, the folding funnel may have dents on the surface representing local energy minima, which can generate kinetic intermediates along the folding pathway. As indicated in the funnel-shaped folding landscape model, it is possible that different unfolded molecules of the same protein may adopt different routes to achieve the same native structure. Depending on which route the protein uses to fold to the native structure, there may or may not be intermediates along the pathway, and there may be preferred pathway(s) existing among different routes (**Figure 1-1**).

In addition to this energetically driven folding landscape, there is another view arguing that proteins fold through specific intermediates along a single predefined folding

pathway. In one version of this protein folding model, proteins appear to fold in units of secondary structures (also termed foldons, cooperative structural units of the native protein).[22] The cooperative foldon units, not the individual amino acids, account for the unit steps in protein folding pathways. A protein appears to first develop the foldons, these building blocks then proceed to grow into global structures as guided by a sequential stabilization mechanism, and this local to global protein folding process that can be fast.[22] The experimental results and predictions by many protein folding studies indicate that the folding process is surprisingly robust that the major character of the folding free-energy landscape is determined by the topology of a protein's native structure. Thus, the cooperative native-like foldon units, together with the sequential stabilization process generate the predetermined stepwise folding pathways.[22]

There are other models proposed to describe the model of protein folding, and proteins may adopt one or more elements from those models in the process of their folding processes. In the nucleation-condensation model, the folding reaction is triggered by the formation of nucleus which is partially stabilized by the formation of some native-like structure and tertiary contacts, followed by additional structure that is put together by rapid condensation using pre-formed nucleus as the template.[23, 24] The diffusion collision model proposes that fluctuating micro-domain structures are formed first, followed by diffusion and collision to form the final structure.[25] Hydrophobic collapse model emphasizes the role of hydrophobic effects, that a rapid collapse of the unfolded peptide chain initiates the formation of the secondary structure.[26]

1.1.3. Cooperativity of protein folding

The existence of preferred folding pathways could be the solution to the Levinthal paradox. Kinetic intermediates used to be viewed as a necessary aspect required to facilitate the protein to find its correct fold. However, many studies have shown that a wide range of small proteins, with fewer than 110 residues, follow a two-state (all-or-none) folding and unfolding, both in a kinetic and an equilibrium sense.[27, 28] Two-state cooperative folding was first implied by calorimetry with the implication of a bimodal conformational distribution with two well-separated states, with a very low population of any intermediate conformations.[29] This can be interpreted as being caused by a very high free energy barrier separating the folded and unfold states, resulting in only two significantly populated states being detected, with no intermediates or partially ordered non-native conformations observed along the folding process.[30] In this case, the surface of the folding funnel model should be a near-Levinthal funnel, featuring a gentle slope on the top which moderates energetic bias to guide the conformation search, and a narrow long stem with a very steep slope, consistent with a very low population of intermediate conformations.[31, 32] For small single domain globular proteins, there are several factors that contribute to the rate of the two-state folding, including the sequence distance of contacts formed in their native structure (topology), local energy distribution, and molecular cooperativity.[27, 28, 33]

U $\rightleftharpoons$ **F** (Two state folding mechanism)

In contrast, intermediate states are often present during the folding of larger proteins, and the folding kinetics of large proteins is mainly determined by escaping from the local minimal free energy state (conformations with low free energy). In addition, the folding of larger proteins is often facilitated by molecular chaperones by preventing protein mis-fold or aggregation.[34]

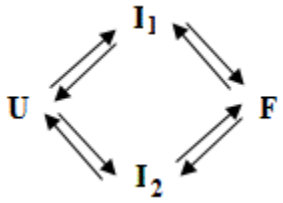
$$U \rightleftharpoons I_1 \rightleftharpoons I_2 \rightleftharpoons I_3 \rightleftharpoons \ldots \rightleftharpoons F$$ (Multi-state folding mechanism)

Furthermore, the intermediates can be classified as "on-pathway" or "off-pathway", depending on the location of the intermediates in the folding scheme.

$$U \rightleftharpoons I \rightleftharpoons F$$ (On-pathway)

$$I \rightleftharpoons U \rightleftharpoons F$$ (Off-pathway)

In addition to the single protein folding pathway, parallel pathways also exist. In this folding scheme, there is more than one transition state between the protein unfolded state and the native state.[35]



(Parallel folding pathway)

1.1.4. Thermodynamics and kinetics of protein folding

The thermodynamic properties of protein folding are important for understanding the stability of proteins. The definition of the protein stability is the difference of the free energy between the protein native state (N) and the unfolded state (U):

$$\Delta G^{o}_{N\text{-}U} = G^{o}_{U} - G^{o}_{N} \qquad \text{(eq 1-1)}$$

In this case, positive values of $\Delta G^{o}_{N\text{-}U}$ indicates that the native state is more stable.

Chemical denaturation and thermal denaturation are the two most commonly used experimental methods to measure the protein stability. The unfolding free energy, ΔG^{o}, often exhibits a linear dependence on the concentration of the denaturant, described by:[36, 37]

$$\Delta G^{\circ}_{U} = \Delta G^{\circ}(H_2O) - m\,[\text{denaturant}] \qquad \text{(eq 1-2)}$$

in which m is a constant relevant to the change of solvent assessable surface area (SASA) in the unfolding process and $\Delta G^{o}(H_2O)$ is the unfolding free energy in the absence of denaturant.[36-38]

For the thermally induced protein unfolding, the unfolding free energy as a function of temperature is defined by the Gibbs-Helmholtz equation:

$$\Delta G^{\circ}(T) = \Delta H^{\circ}(T_m)\,(1 - T/T_m) + \Delta C^{\circ}_{P}\,[(T - T_m) - T\ln(T/T_m)] \qquad \text{(eq 1-3)}$$

in which T_m is the thermally induced unfolding midpoint temperature, the point with equal populations of the folded and unfolded protein ensembles present (**Figure 1-2**), $\Delta H^\circ(T_m)$ is the enthalpy change at T_m, and ΔC°_P is the heat capacity change upon unfolding at T_m with the assumption that ΔC°_P is independent of the temperature. ΔC°_P, similar to *m*, is also related to the change of SASA upon unfolding, and is often thought to be dominated by the change in the hydrophobic surface area between the native state and the unfolded state.[39]

The Gibbs-Helmholtz equation also predicts that proteins will unfold as the temperature is decreased. This process, referred to as cold denaturation, is described in **section 1.3**.

Kinetic measurements of the protein folding process provides information on the mechanism of protein folding. As described earlier, the folding of a protein can range from microseconds to seconds, depending on the size and the amino acid sequence pattern. The protein folding rate k_f and unfolding rate k_u depend on the concentration of denaturant and follow the equation below for a two-state folder:

$$\ln(k_{obs}) = \ln(k_{f(H_2O)}e^{m_f[den]/RT} + k_{u(H_2O)}e^{m_u[den]/RT}) \qquad \text{(eq 1-4)}$$

where k_f (H_2O) and k_u (H_2O) are the folding and unfolding rate in the absence of denaturant, [den] is the concentration of the denaturant, m_f and m_u are constants defined by the change of $\ln(k_f)$ and $\ln(k_u)$ as a function of the [den]. The parameter m_f is related to the change of SASA between the folded state and the transition state, and is negative,

while the parameter m_u is related to the change of SASA between the unfolded state and the transition state, and is positive. The equilibrium *m*-value, m_{total}, is the difference between m_u/RT and m_f/RT. In a two-state folding process, a plot of k_{obs} against [den] is a V-shaped curve, also termed as Chevron plot.[40]

1.2. The protein denatured state ensemble (DSE) or protein unfolded state

In order to get a full picture of the protein folding pathway, a detailed characterization of all the species along the folding pathway is required, including the folded state, the unfolded state, and any folding intermediates, and in so far as possible, the transition states. Characterization of native folded structure of many proteins by X-ray crystallography and nuclear magnetic resonance (NMR) has been well-established for many years.[41-44] The transition states for the two-state folding reaction have been investigated by ϕ-value analysis. The ϕ-value analysis is the ratio of the change in the activation free energy due to the mutation over the change in the equilibrium free energy caused by the same mutation.[45, 46] The ϕ-value can be interpreted as an indication of how well a native interaction is developed in the transition state, based on the assumption that the mutation does not perturb the unfolded state.[47]

Relatively little is known for the protein unfolded state, especially unfolded states populated under near native conditions.

1.2.1. The significance of studies on the protein unfolded state

Studies on the characterization of the protein unfolded state are important for understanding the mechanism of the protein folding process. The protein unfolded state, or the protein denatured state ensemble (DSE), is the starting point of the protein folding process and the DSE is the thermodynamic reference state for all studies of protein stability.[48-50] In addition, the DSE can be a target for rational protein design and engineering, since the overall stability of a designed protein will be enhanced by destabilizing the DSE.[51, 52] Information on factors that impact protein misfolding, the tendency for protein aggregation *in vitro* and *in vivo,* and amyloid formation can be obtained from the studies of the properties of the DSE.[53-55] From studies during the past decade, it has been found that many functional proteins do not have well-folded structures, and these so-called intrinsically disordered proteins (IDPs) and proteins with intrinsically disordered regions (IDRs) are highly abundant in living organisms and play vital roles *in vivo*.[56-58] The exploration of the mechanisms and biological function of the IDPs largely depends on the characterization of the properties of unstructured and partially structured states and therefore has much in common with studies of the DSE.[57, 59]

1.2.2. The random coil model is not accurate to describe the protein DSE

Unlike the well-defined structure that the protein native state has, the protein unfolded state contains a number of rapidly inter-converting conformations. Traditionally, the protein unfolded state was considered to be a random coil, adapted from polymer science,

which was developed by Flory and corroborated by Tanford.[60-62] A random coil is a polymer chain in which every single bond is moving freely and without any specific contacts or preferred conformations. By definition, the random coil model is an energetically well-defined reference state in which no side-chain side-chain interactions are present.[50]

While some proteins appear to behave as random coils under strongly denaturing conditions, there is excellent evidence that unfolded states under more native like conditions often do not. Many recent experimental data suggest the protein denatured state under some conditions can be highly compact and may contain native and/or non-native secondary structure.[63-68] Residual native structure in the protein unfolded state can reduce the initial conformational search and accelerate protein folding.[69]

1.2.3. The structure of the protein DSE depends on experimental conditions

There are multiple ways to denature a folded protein. The protein DSE populated under strongly denaturing conditions, such as at extreme pH or under high concentration of denaturant, is one method of studying the structural properties of the DSE. It is easy to populate the protein DSE under those conditions experimentally and the structural properties of the DSE can be studied at equilibrium. Chemical denaturation promotes the unfolding of the proteins through more extensive favorable association between the denaturant and the unfolded protein.[50] Additional protons attached to the residues at low pH and the destabilization of the native structure caused by the resulting charge repulsion

can be the explanation for the acid-induced protein unfolding. Some proteins can also be unfolded at high pH and the effect is also likely due to charge repulsion. Under the harsh denaturing conditions mentioned above, the protein DSE is mostly expanded as judged by global parameters such as the radius of hydration (R_h) and the radius of gyration (R_g). The global parameter R_g scales with the chain length, N, as $N^{0.59}$, for proteins unfolded in denaturant which is in accordance to the random coil model. This scaling behavior was examined for 17 chemically denatured proteins by small angle X-ray scattering (SAXS).[70] DSEs having similar R_g values are normally considered to be similarly unfolded, but this has not rigorously tested.[70, 71] In apparent contrast, several studies of protein DSEs populated under strongly denaturing conditions reveal the presence of residual structure, which deviates from the concept of random coil model. For example, the N-terminal domain of the 434-repressor was found to have a hydrophobic cluster formed in 7 M urea by Nuclear Overhauser effect (NOE) measurements.[65] Significant deviations from the random coil model have been reported on intact lysozyme in urea.[72, 73] The residual structure formed in these denatured proteins was suggested to be important in the early stage of the protein folding process. The interactions that cause the polypeptide chain to deviate from random coil model are under debate. Under conditions such as low pH and high denaturant concentration, it has been reported that the residual structure is due to the formation of aromatic interactions.[65, 68, 74-78] Clusters in the DSE are stabilized by both native and non-native long-range hydrophobic interactions, as demonstrated by further analysis on the nature of the clusters.[76]

Despite the disagreement on the extent of residual structure that can be formed in the protein DSE generated by chemical denaturant, the mutual existence of residual structure and random coil behavior has been reported.[70, 79, 80] More recent work has shown that residual unfolded state structure is compatible with the $N^{0.59}$ scaling law. Conformational ensembles generated from models with native like contacts still exhibit random coil scaling as shown by R_g values comparable to those measured experimentally.[81-84]

The DSE populated under near native conditions which co-exists with the folded state is the state that is directly relevant to protein folding under native conditions. It is very difficult to populate both the folded and unfolded state simultaneously, because of the cooperativity of the protein folding and because the free energy balance usually strongly favors the folded state under native conditions.[85, 86] In favorable cases, such as the drk SH3 domain, mild conditions can be found to study the unfolded state in equilibrium with the folded ensembles.[64, 87, 88] Mutants with lower stability compared to the wild type protein can sometimes be used to populating the DSE in native buffer.[89] Currently, there is an ongoing debate regarding whether the DSE under near native state is more compact than the DSE populated under strongly denatured conditions.[86] The argument arises due to the different results obtained from SAXS and single-molecule FRET (smFRET) studies. For example, for the small 64-residue protein, protein L, SAXS studies showed that the R_g under low concentration of denaturant is as expanded as that in high concentration of denaturant, whereas smFRET showed a significant compaction of R_g for the DSE populated under low denaturant conditions.[86] Despite the debates on the dimensional compactness of DSE under near native conditions, many NMR studies show

that more structure is formed in the DSE populated under near native conditions, compared with the strongly denatured conditions.[85]

1.2.4. Methods to study the protein denatured state ensemble

One way to directly compare the differences between the protein DSE under native conditions and the DSE under different conditions is to examine the overall dimensions or compactness. Pulse Field Gradient Nuclear Magnetic Resonance (PFG-NMR) and SAXS are frequently used to measure R_h and R_g respectively, thus allowing direct comparison of DSE under different conditions.[71, 90-92] But these studies do not reveal detailed differences, such as long-range interactions or local structure that may exist within the DSE. Circular dichroism (CD) and fluorescence are both sensitive to the protein structure, and they are frequently used to monitor the protein folding process, providing thermodynamics parameters, and probing overall secondary structure or burial of residues.

In order to obtain residue-specific information at an atomic level, NMR has great advantages. With the development of isotopically labeled proteins and multidimensional heteronuclear pulse sequences, backbone and side-chain assignment can be achieved, allowing tracking of changes on a specific residue/atom.[93-97] Valuable information on the dynamic and structural properties of the protein DSE can be obtained using parameters such as chemical shifts, Nuclear Overhauser Effects (NOEs), relaxation rate constants,

and residual dipolar couplings (RDCs).[98-103] This atomic level information together with the global parameters will provide a more detailed description of the protein DSE.

NOEs are used to detect short range interactions with in 5 Å, are often utilized to detect helical or turn-like structure in the protein DSE.[64, 88, 104] As for long range interactions, paramagnetic relaxation enhancement (PRE)-NMR experiments are widely used to identify average distance up to 30 Å in a dynamic structure, such as in the protein DSE studies.[67, 105-111] PRE depends on the magnetic interaction between the unpaired electron in the nitroxide spin label (which is covalently attached to a Cys side-chain) and a proton (usually the amide proton). This interaction increases the transverse relaxation rate of the proton by a factor that is proportional to the r^{-6}, in which r is the distance between the two interacting protons, thus the transiently formed long range interactions in the DSE can be picked up by the PRE measurements.[105, 106] The long range interactions in the partially unfolded or the unfolded proteins can reveal how native order initiates and propagates from the earliest stage of protein folding process. The relationship between the long-range interactions and the hydrophobic interactions have been examined.[107, 112, 113]

1.3. Protein cold denaturation

Protein cold denaturation, the transition from the folded state to an unfolded state induced by lowering the temperature from the temperature of maximal stability (T_s or T_{max}) [114], is a general property of globular proteins. Cold denaturation for β-lactoglobulin under

equilibrium conditions was first reported in 1948. [115] Protein cold denaturation is a general feature that well predicted by Gibbs-Helmholtz equation, as the temperature drops the enthalpy of folding decreases and eventually the denatured state is favored (**Figure 1-2**). Hydrophobic effects have been suggested to induce the cold denaturation of globular proteins: as the temperature decreases, the free energy penalty of the entropically unfavorable interactions between hydrophobs and water becomes smaller.[116-125] Protein cold denaturation is also very important since it is relevant to the formulation, freeze-thaw and storage conditions of the therapeutic proteins.[126-128] Evidence for protein denaturation *in vivo* has been found in cold-adapted organisms.[129-131] A more detailed understanding of protein cold denaturation is of considerable practical importance. The midpoint of cold denaturation, T_c, can be derived from the Gibbs-Helmholtz equation and described as: [132]

$$T_c = \frac{{T_m}^2}{T_m + 2(\Delta H^o(T_m)/\Delta {C_p}^o)} \qquad \text{(eq 1-5)}$$

T_m is the midpoint temperature of the thermally induced unfolding, $\Delta H^\circ(T_m)$ is the enthalpy change at T_m, and ΔC°_P is the heat capacity change upon unfolding at T_m with the assumption that ΔC°_P is independent of the temperature. High values of T_c is the result of the ratio of $\Delta H^\circ(T_m)/ \Delta C^\circ_P$ is small.

The midpoint of protein cold denaturation, T_c, is commonly well below the freezing point of water, preventing studies of protein cold denaturation in aqueous solution. Modifications of the system such as adding denaturant, high pressure, extremes of pH, or

reverse micelle encapsulation have allowed studies of protein cold denaturation.[114, 125, 133-139] But those conditions are different from native aqueous buffer, and proteins may behave quite differently under those conditions. Unlike the well accepted two-state, cooperative thermal and denaturant induced unfolding for small single-domain proteins, the cooperativity of cold denaturation is still controversial.[140-148] The yeast orthologue of frataxin (Yfh1), a small α-β protein undergoes two-state cold unfolding.[146-148] However, deviations from the two-state cold denaturation and a step-wise folding scheme has been proposed for ubiquitin carried out by reverse micelle encapsulation studies.[141, 142, 145] But temperature-dependent interactions between the reverse micelles and the proteins being encapsulated, together with water shedding may complicate the system under observation, thus the nature of the cold denaturation may be biased in such systems.[144, 149] To gain more insight into the cooperativity of the protein cold denaturation process, it is desirable to study systems under near physiological conditions.

1.4. The C-terminal domain of the ribosomal protein L9

The structural characterization of the protein unfolded state described in this dissertation was carried out using the model protein, the C-terminal domain of the ribosomal protein L9 (CTL9). CTL9 is an isolated domain of the intact full-length, 149-residue ribosomal protein L9 from the bacteria *Geobacillus stearothermophilus* (*Bacillus stearothermophilus*, which has been reclassified to be a member of the genus *Geobacillus* in 2001), a thermophile that is widely found in ocean sediment, soil, and hot springs. The C-terminal domain consists of the last 92 residues of L9.

1.4.1. The structure of the intact ribosomal protein L9

The intact ribosomal protein L9 (**Figure 1-3**), a 149-residue protein, is composed of two individual globular domains: the N-terminal domain and the C-terminal domain, denoted as NTL9 and CTL9 respectively. The two domains are connected by a central solvent exposed helix linker, which serves to fix the relative domain separation and orientation within the ribosome.[150] L9 does not form any disulfide bonds, and it does not bind to any metal ions or cofactors. Both of the two domains individually follow a two-state folding pathway and exhibit independent folding processes.[151, 152] The C-terminal domain folds at a significantly slower rate compared to the N-terminal domain, likely because of the mixed parallel, antiparallel β-sheet within the domain.[152] L9 has a high yield when expressed in *E. coli* and it is highly soluble. Thus it is an excellent model system.

1.4.2. Biological functions of ribosomal protein L9

It is extremely difficult to define the role of a particular ribosomal protein in protein synthesis, but based on the location of the protein L9 in the ribosomal protein complex, the function in general terms has been determined. From the X-ray crystal structure of the bacterial ribosome, the location of L9 in the ribosome was revealed to be in the large subunit 50S and close to protein L2 and L28.[153, 154] Each domain of the protein L9 binds to the domain V region of 23S rRNA, and is thought to act as a molecular scaffold to

stabilize the ribosome.[150, 155-157] However, the biological function of this protein is not fully understood yet.

1.4.3. The isolated C-terminal domain of the ribosomal protein L9

The isolated C-terminal domain (CTL9, residues 58-149 from the intact L9) has an unusual mixed parallel, antiparallel three-stranded β sheet structure, together with two loops with partial α-helix 3_{10} turns, and two α-helices. The β sheet packs against the central α-helix, as shown in the ribbon diagram (**Figure 1-4**). CTL9 has a high yield of more than 80 mg/L when expressed in *E. coli* and it has high solubility in water.

Previous studies in our group have shown that CTL9 follows a two-state folding model and can be denatured by lowering the pH, by adding denaturant, or at high/low temperature. The three His in the primary sequence makes the folding of CTL9 strongly pH-dependent and reversible unfolding has been observed in the pH range from 10 to 2.0, in the absence of denaturant, at room temperature. At pH values higher than 5.0, CTL9 is fully folded, while at pH values below 2.5, the protein is populated as the acid-induced CTL9 DSE (the acid unfolded state). Both the folded and unfolded forms of CTL9 coexist between pH 2.5 and 5.0, allowing the study of the structure of the acid-induced CTL9 DSE and the folded state under identical solvent conditions.[158] The expansion of the acid-induced CTL9 DSE at lower pH is consistent with the results of thermodynamic studies, which reported an increase in both the *m*-values and the heat capacity as the pH decreases.[159] Hydrodynamic properties, probed using pulse field gradient (PFG)-diffusion

measurements, revealed that the acid-induced CTL9 DSE at pD 2.1 is expanded compared with the folded ensemble, and it has a very similar radius of hydration (R_h) to that of the 8 M urea-induced DSE.[160] In spite of the similar values of R_h and seemingly the same random coil CD spectra, the acid-induced DSE exhibits a more significant propensity to form secondary structure than the urea-induced DSE, as suggested by NMR and secondary structure prediction (SSP) analysis. SSP analysis uses a weighted average of the secondary chemical shits to predict the tendency of forming secondary structure.[161] In addition, the tendency to adopt both native and non-native helical structure in the DSE populated in the absence of denaturant is stronger at pH 3.8, indicating that the DSE is more structured under native-like conditions, compared to the fully unfolded states.[85]

Another interesting feature of CTL9 is that one point mutant, I98A-CTL9, alters the core packing of the protein and significantly destabilizes the protein stability by 4 kcal/mol. This mutant allows us to study the protein cold denaturation in water above 0 °C, without system modifications or perturbations. The cold denatured state of CTL9 is compact, containing both native and non-native structure, but it has been shown to become more expanded as the temperature decreased from 25 to 2 °C, and a 20% increase in the R_h has been observed.[162, 163]

1.5. The aims of this dissertation

The aims of the research described in my thesis are to characterize the structural and dynamic properties of the DSE of CTL9 populated under various denaturing conditions,

to study the folding cooperativity of I98A-CTL9 cold denaturation, and to compare the properties of temperature dependent I98A-CTL9 DSE to the cold unfolded DSE. NMR was used to provide residue-specific information at the atomic level. In addition, SAXS was used to monitor the overall shape and compactness. The acid- and urea-induced DSE of CTL9 showed significantly different long-range contacts within the unfolded ensemble, despite the fact that they have similar radius of gyration (R_g) (**Chapter Two**). The properties of the DSE are populated in equilibrium with the folded state were examined as a function of pH and urea concentration, respectively (**Chapter Three**). The cooperativity of cold denaturation was examined and its two-state nature was confirmed using the core destabilizing point mutant I98A-CTL9 as a model system (**Chapter Four**). Finally, the properties of cold- and heat-induced I98A-CTL9 DSEs were compared in terms of their compactness and their tendency to form secondary structure (**Chapter Five**).

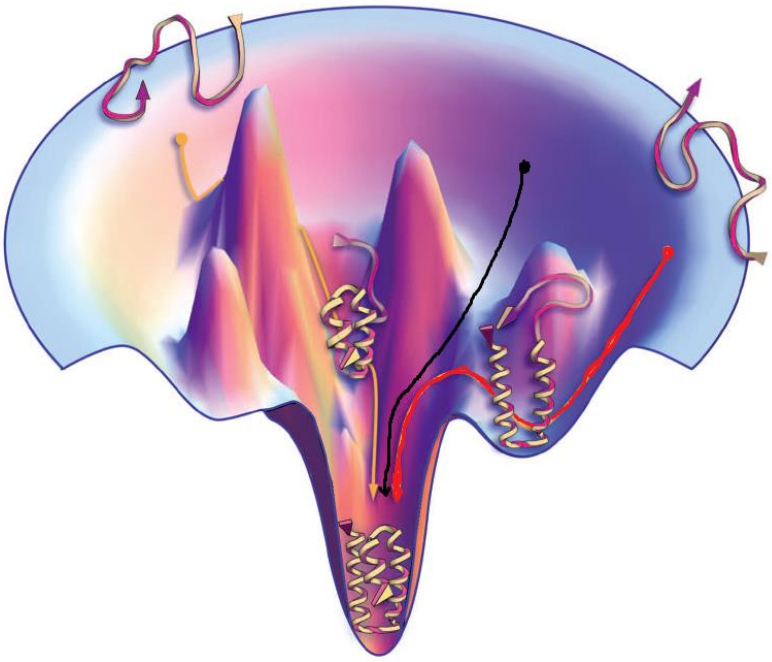


Figure 1-1. The funnel shaped free energy landscape folding model with three different folding pathways shown in black, yellow and red. The unfolded protein ensembles with a wide range of conformations undergo the folding process to form the single unique native folded structure, the folding is driven by the energetically downhill nature of the free energy surface. Three different folding pathways are shown, (1) Black, no intermediates formed in the folding pathway, (2) Yellow and (3) Red: they have different intermediate states in the middle of the folding process. The figure is adapted and modified from ref [3]. The vertical axis represents the total free energy, G_T, minus the contribution due to configuration entropy times temperature, $-TS_{config}$, i.e. $G_T - TS_{config}$. The width of the funnel is proportional to the conformational entropy.

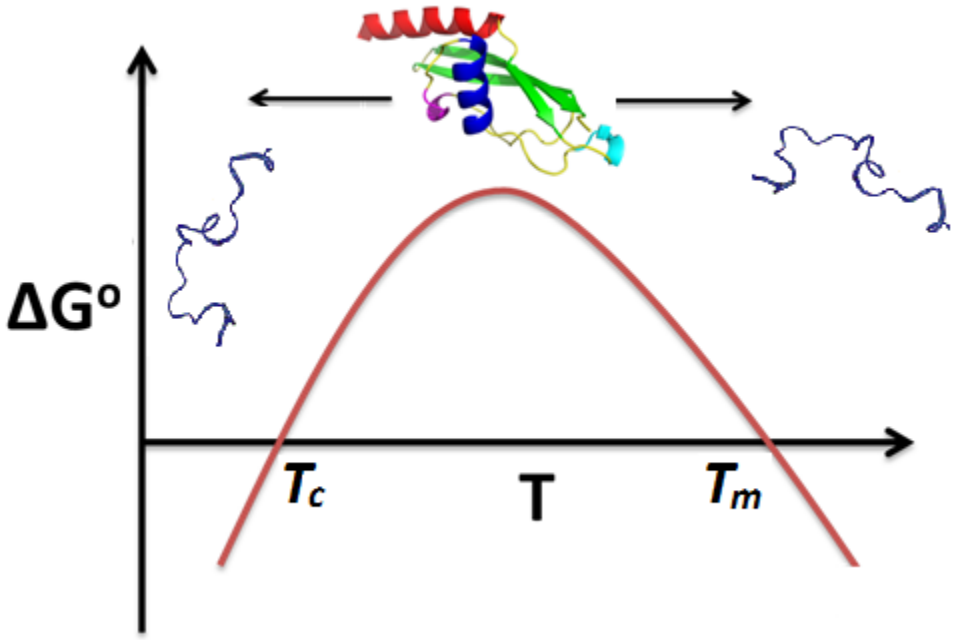


Figure 1-2. A schematic plot of the free energy change of the protein unfolding reaction as a function of temperature, as described by Gibbs-Helmholtz equation. The transition temperature of the protein cold denaturation (T_c) and the transition temperature (T_m) of the protein heat or thermal denaturation are shown. The transition of maximum stability, T_s, occurs when $\partial\Delta G^o/\partial T = -\Delta S^o = 0$. So T_s is also the temperature at which the net entropy change is zero.

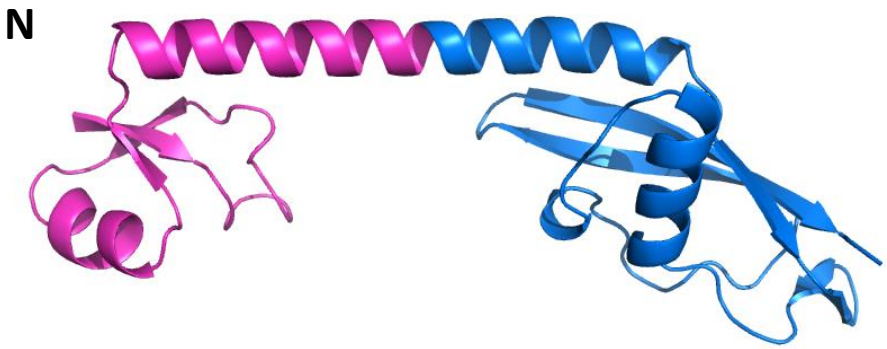


Figure 1-3. Structure of the ribosomal protein L9, shown as ribbon diagram using the PDB file 1DIV. The N-terminus is labeled and the region corresponding to the C-terminal domain construct, denoted CTL9, is shown in blue.

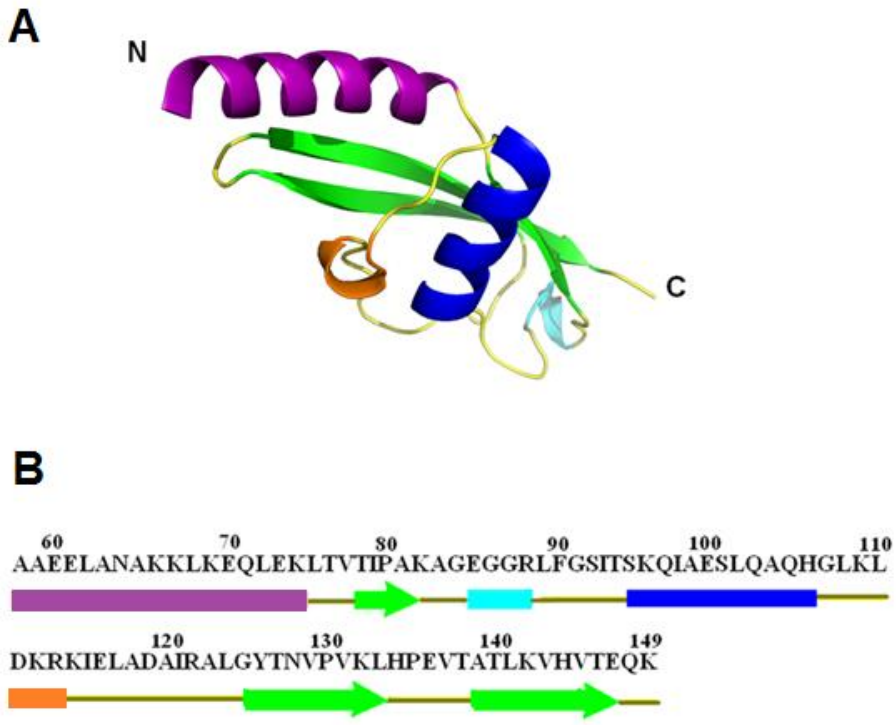


Figure 1-4. (A) Structure of the C-terminal domain of the ribosomal protein L9 (CTL9), shown as ribbon diagram, PDB file 1DIV. The N- and C-terminus are labeled. (B) The primary sequence of CTL9 is shown below with the matched color-coding.

1.6. References:

[1] Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900-904.

[2] Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annu Rev Biochem 75*, 333-366.

[3] Eisenberg, D., Nelson, R., Sawaya, M. R., Balbirnie, M., Sambashivan, S., Ivanova, M. I., Madsen, A. O., and Riekel, C. (2006) The structural biology of protein aggregation diseases: Fundamental questions and some answers, *Acc Chem Res 39*, 568-575.

[4] Ma, Y., and Hendershot, L. M. (2004) The role of the unfolded protein response in tumour development: Friend or foe?, *Nat Rev Cancer 4*, 966-977.

[5] Onuchic, J. N., and Wolynes, P. G. (2004) Theory of protein folding, *Curr Opin Struct Biol 14*, 70-75.

[6] Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on, *Science 338*, 1042-1046.

[7] Moult, J. (2005) A decade of CASP: Progress, bottlenecks and prognosis in protein structure prediction, *Curr Opin Struct Biol 15*, 285-289.

[8] Bradley, P., Misura, K. M., and Baker, D. (2005) Toward high-resolution de novo structure prediction for small proteins, *Science 309*, 1868-1871.

[9] Zhou, H. X., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences, *Annu Rev Biophysics 37*, 375-397.

[10] Cheung, M. S., Klimov, D., and Thirumalai, D. (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins, *Proc Natl Acad Sci USA 102*, 4753-4758.

[11] Horwich, A. L., Farr, G. W., and Fenton, W. A. (2006) GroEL-GroES-mediated protein folding, *Chem Rev 106*, 1917-1930.

[12] Anfinsen, C. B. (1972) The formation and stabilization of protein structure, *Biochem J 128*, 737-749.

[13] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223-230.

[14] Hart, T., Hosszu, L. L., 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.

[15] 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.

[16] 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.

[17] Levintha.C. (1968) Are there pathways for protein folding, *J Chim Phys 65*, 44-45.

[18] Dobson, C. M., and Karplus, M. (1999) The fundamentals of protein folding: Bringing together theory and experiment, *Curr Opin Struct Biol 9*, 92-101.

[19] Dobson, C. M., Šali, A., and Karplus, M. (1998) Protein Folding: A perspective from theory and experiment, *Angew Chem Int Ed 37*, 868-893.

[20] Chan, H. S., and Dill, K. A. (1998) Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics, *Proteins 30*, 2-33.

[21] 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.

[22] Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: Mechanism and principles, *Q Rev Biophys 40*, 287-326.

[23] Fersht, A. R. (1997) Nucleation mechanisms in protein folding, *Curr Opin Struct Biol 7*, 3-9.

[24] 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.

[25] Karplus, M., and Weaver, D. L. (1994) Protein folding dynamics: The diffusion-collision model and experimental data, *Protein Sci 3*, 650-668.

[26] Dill, K. A. (1985) Theory for the folding and stability of globular proteins, *Biochemistry 24*, 1501-1509.

[27] Barrick, D. (2009) What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding?, *Phys Biol 6*.

[28] Baker, D. (2000) A surprising simplicity to protein folding, *Nature 405*, 39-42.

[29] Lumry, R., and Biltonen, R. (1966) Validity of the "two-state" hypothesis for conformational transitions of proteins, *Biopolymers 4*, 917-944.

[30] Kaya, H., and Chan, H. S. (2000) Polymer principles of protein calorimetric two-state cooperativity, *Proteins 40*, 637-661.

[31] Kaya, H., and Chan, H. S. (2003) Simple two-state protein folding kinetics requires near-levinthal thermodynamic cooperativity, *Proteins 52*, 510-523.

[32] Zwanzig, R., Szabo, A., and Bagchi, B. (1992) Levinthal's paradox, *Proc Natl Acad Sci USA 89*, 20-22.

[33] Jackson, S. E. (1998) How do small single-domain proteins fold?, *Fold Des 3*, R81-R91.

[34] Horwich, A. L., Weber-Ban, E. U., and Finley, D. (1999) Chaperone rings in protein folding and degradation, *Proc Natl Acad Sci USA 96*, 11033-11040.

[35] Fersht, A. R., Itzhaki, L. S., elMasry, N. F., Matthews, J. M., and Otzen, D. E. (1994) Single versus parallel pathways of protein folding and fractional formation of structure in the transition state, *Proc Natl Acad Sci USA 91*, 10426-10429.

[36] Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol 131*, 266-280.

[37] Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Adv Protein Chem 24*, 1-95.

[38] Scholtz, J. M., Grimsley, G. R., and Pace, C. N. (2009) Solvent denaturation of proteins and interpretations of the m value, *Methods Enzymol 466*, 549-565.

[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] 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.

[41] Clore, G. M., and Gronenborn, A. M. (1998) Determining the structures of large proteins and protein complexes by NMR, *Trends Biotechnol 16*, 22-34.

[42] Clore, G. M., and Gronenborn, A. M. (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa, *Curr Opin Chem Biol 2*, 564-570.

[43] Wagner, G. (1997) An account of NMR in structural biology, *Nature Struct Biol 4 Suppl*, 841-844.

[44] Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Evdokimova, E., Liu, C. S., Semesi, A., Guido, V., Edwards, A. M., and Arrowsmith, C. H. (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins, *J Am Chem Soc 127*, 16512-16517.

[45] Daggett, V., and Fersht, A. (2003) The present view of the mechanism of protein folding, *Nat Rev Mol Cell Biol 4*, 497-502.

[46] Dalby, P. A., Oliveberg, M., and Fersht, A. R. (1998) Folding intermediates of wild-type and mutants of barnase. I. Use of ϕ-value analysis and m-values to probe the cooperative nature of the folding pre-equilibrium, *J Mol Biol 276*, 625-646.

[47] Cho, J. H., and Raleigh, D. P. (2006) Denatured state effects and the origin of nonclassical ϕ values in protein folding, *J Am Chem Soc 128*, 16492-16493.

[48] Bowler, B. E. (2007) Thermodynamics of protein denatured states, *Mol BioSys 3*, 88-99.

[49] Baldwin, R. L. (2002) A new perspective on unfolded proteins, *Adv Protein Chem 62*, 361-367.

[50] Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability, *FASEB J 10*, 27-34.

[51] 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.

[52] Anil, B., Song, B., Tang, Y., and Raleigh, D. P. (2004) Exploiting the right side of the Ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability, *J Am Chem Soc 126*, 13194-13195.

[53] Pavitt, G. D., and Ron, D. (2012) New insights into translational regulation in the endoplasmic reticulum unfolded protein response, *Cold Spring Harbor Perspect Biol 4*, DOI: 10.1101/cshperspect.a012278.

[54] Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways, *Arch Biochem Biophys 469*, 100-117.

[55] Polverino de Laureto, 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.

[56] Uversky, V. N. (2013) Unusual biophysics of intrinsically disordered proteins, *Biochim Biophys Acta 1834*, 932-951.

[57] Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nature Rev Mol Cell Biol 6*, 197-208.

[58] Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *J Mol Graphics Modell 19*, 26-59.

[59] Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics, *Protein Sci 11*, 739-756.

[60] Tanford, C. (1968) Protein denaturation, *Adv Protein Chem 23*, 121-282.

[61] Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*, Wiley, New York.

[62] Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Itchaca, NY.

[63] 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.

[64] 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.

[65] 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.

[66] Sung, Y. H., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family, *J Mol Biol 372*, 689-707.

[67] 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.

[68] 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.

[69] Wong, K. B., Freund, S. M., 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.

[70] 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., Hasan, 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.

[71] Millett, 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 62*, 241-262.

[72] Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., and Dobson, C. M. (1996) Toward a description of the conformations of denatured states of proteins. Comparison of a random coil model with NMR measurements, *J Phys Chem 100*, 2661-2666.

[73] Bolin, K. A., Pitkeathly, M., Miranker, A., Smith, L. J., and Dobson, C. M. (1996) Insight into a random coil conformation and an isolated helix: Structural and dynamical characterisation of the C-helix peptide from hen lysozyme, *J Mol Biol 261*, 443-453.

[74] Ropson, I. J., and Frieden, C. (1992) Dynamic NMR spectral analysis and protein folding: Identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR, *Proc Natl Acad Sci USA 89*, 7222-7226.

[75] Saab-Rincon, G., Gualfetti, P. J., and Matthews, C. R. (1996) Mutagenic and thermodynamic analyses of residual structure in the α subunit of tryptophan synthase, *Biochemistry 35*, 1988-1994.

[76] 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, *Angewan Chem 43*, 5780-5785.

[77] Schlorb, C., Mensch, S., Richter, C., and Schwalbe, H. (2006) Photo-CIDNP reveals differences in compaction of non-native states of lysozyme, *J Am Chem Soc 128*, 1802-1803.

[78] Schlorb, C., Ackermann, K., Richter, C., Wirmer, J., and Schwalbe, H. (2005) Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states, *J Biomol NMR 33*, 95-104.

[79] Yi, Q., Scalley-Kim, M. L., Alm, E. J., and Baker, D. (2000) NMR characterization of residual structure in the denatured state of protein L, *J Mol Biol 299*, 1341-1351.

[80] McCarney, E. R., Kohn, J. E., and Plaxco, K. W. (2005) Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins, *Crit Rev Biochem Mol Biol 40*, 181-189.

[81] Meng, W., Lyle, N., Luan, B., Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure, *Proc Natl Acad Sci USA 110*, 2123-2128.

[82] Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc Natl Acad Sci USA 101*, 12497-12502.

[83] Wang, Z., Plaxco, K. W., and Makarov, D. E. (2007) Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins, *Biopolymers 86*, 321-328.

[84] Voelz, V. A., Singh, V. R., Wedemeyer, W. J., Lapidus, L. J., and Pande, V. S. (2010) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment, *J Am Chem Soc 132*, 4702-4709.

[85] 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.

[86] Yoo, T. Y., Meisburger, S. P., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., and Plaxco, K. (2012) Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state, *J Mol Biol 418*, 226-236.

[87] Zhang, O. W., and FormanKay, J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions, *Biochemistry 36*, 3959-3970.

[88] Zhang, O., and Formankay, J. D. (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, *Biochemistry 34*, 6784-6794.

[89] Meng, W., Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2013) The denatured state ensemble contains significant local and long-range structure under native conditions: Analysis of the N-terminal domain of ribosomal protein L9, *Biochemistry 52*, 2662-2671.

[90] Jones, J., Wilkins, D., Smith, L., and Dobson, C. (1997) Characterisation of protein unfolding by NMR diffusion measurements, *J Biomol NMR 10*, 199-203.

[91] 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.

[92] Choy, W. Y., Mulder, F. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and Kay, L. E. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques, *J Mol Biol 316*, 101-112.

[93] Eliezer, D. (2007) Characterizing residual structure in disordered protein States using nuclear magnetic resonance, *Methods Mol Biol 350*, 49-67.

[94] Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR, *Chem Rev 104*, 3607-3622.

[95] Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J Biomol NMR 20*, 71-75.

[96] Fersht, A. R., and Daggett, V. (2002) Protein folding and unfolding at atomic resolution, *Cell 108*, 573-582.

[97] Dyson, H. J., and Wright, P. E. (1998) Equilibrium NMR studies of unfolded and partially folded proteins, *Nat Struct Biol 5 Suppl*, 499-503.

[98] 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 222*, 311-333.

[99] 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.

[100] 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.

[101] Ohnishi, S., and Shortle, D. (2003) Observation of residual dipolar couplings in short peptides, *Proteins 50*, 546-551.

[102] Fieber, W., Kristjansdottir, S., and Poulsen, F. M. (2004) Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings, *J Mol Biol 339*, 1191-1199.

[103] Tollinger, M., Skrynnikov, N. R., Mulder, F. 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.

[104] Tang, Y., Goger, M. J., and Raleigh, D. P. (2006) NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain, *Biochemistry 45*, 6940-6946.

[105] 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.

[106] 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.

[107] 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.

[108] Thomsen, J. K., Kragelund, B. B., Teilum, K., Knudsen, J., and Poulsen, F. M. (2002) Transient intermediary states with high and low folding probabilities in the apparent two-state folding equilibrium of ACBP at low pH, *J Mol Biol 318*, 805-814.

[109] Lindorff-Larsen, 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.

[110] Xue, Y., Podkorytov, I. S., Rao, D. K., Benjamin, N., Sun, H., and Skrynnikov, N. R. (2009) Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain, *Protein Sci 18*, 1401-1424.

[111] Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, *J Mol Biol 324*, 349-357.

[112] Vendruscolo, M., and Paci, E. (2003) Protein folding: Bringing theory and experiment closer together, *Curr Opin Struc Biol 13*, 82-87.

[113] Dyson, H. J., Wright, P. E., and Scheraga, H. A. (2006) The role of hydrophobic interactions in initiation and propagation of protein folding, *Proc Natl Acad Sci USA 103*, 13057-13061.

[114] Brandts, J. F. (1964) Thermodynamics of protein denaturation .I. Denaturation of chymotrypsinogen, *J Am Chem Soc 86*, 4291-4292.

[115] Jacobsen, C. F., and Christensen, L. K. (1948) Influence of temperature on the urea denaturation of β-lactoglobulin, *Nature 161*, 30-31.

[116] Dias, C. L. (2012) Unifying microscopic mechanism for pressure and cold denaturations of proteins, *Phys Rev Lett 109*, 048104.

[117] Graziano, G. (2010) On the molecular origin of cold denaturation of globular proteins, *Phys Chem Chem Phys 12*, 14245-14252.

[118] Yoshidome, T., and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein, *Phys Rev E 79*, 030905.

[119] 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.

[120] Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I., and Grant, M. (2008) Microscopic mechanism for cold denaturation, *Phys Rev Lett 100*, 118101.

[121] Dias, C. L., Ala-Nissila, T., Wong-ekkabut, J., Vattulainen, I., Grant, M., and Karttunen, M. (2010) The hydrophobic effect and its role in cold denaturation, *Cryobiology 60*, 91-99.

[122] Lopez, C. F., Darst, R. K., and Rossky, P. J. (2008) Mechanistic elements of protein cold denaturation, *J Phys Chem B 112*, 5961-5967.

[123] Tsai, C. J., Maizel, J. V., Jr., and Nussinov, R. (2002) The hydrophobic effect: A new insight from cold denaturation and a two-state water structure, *Crit Rev Biochem Mol Biol 37*, 55-69.

[124] Franks, F. (1995) Protein destabilization at low temperatures, *Adv Protein Chem 46*, 105-139.

[125] Privalov, P. L., Griko Yu, V., Venyaminov, S., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin, *J Mol Biol 190*, 487-498.

[126] Lazar, K. L., Patapoff, T. W., and Sharma, V. K. (2010) Cold denaturation of monoclonal antibodies, *mAbs 2*, 42-52.

[127] Feller, G. (2010) Protein stability and enzyme activity at extreme biological temperatures, *J Phys 22*, 323101.

[128] Kolhe, P., and Badkar, A. (2011) Protein and solute distribution in drug substance containers during frozen storage and post-thawing: A tool to understand and define freezing-thawing parameters in biotechnology process development, *Biotechnol Progr 27*, 494-504.

[129] Hayley, M., Chevaldina, T., and Heeley, D. H. (2011) Cold adaptation of tropomyosin, *Biochemistry 50*, 6559-6566.

[130] Todgham, A. E., Hoaglund, E. A., and Hofmann, G. E. (2007) Is cold the new hot? Elevated ubiquitin-conjugated protein levels in tissues of Antarctic fish as evidence for cold-denaturation of proteins *in vivo*, *J Comp Phys B 177*, 857-866.

[131] Pucciarelli, S., Parker, S. K., Detrich, H. W., 3rd, and Melki, R. (2006) Characterization of the cytoplasmic chaperonin containing TCP-1 from the Antarctic fish *Notothenia coriiceps*, *Extremophiles 10*, 537-549.

[132] Privalov, P. L. (1990) Cold denaturation of proteins, *Crit Rev Biochem Mol Biol 25*, 281-305.

[133] Tamura, A., Kimura, K., Takahara, H., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .1. CD and DSC studies, *Biochemistry 30*, 11307-11313.

[134] Tamura, A., Kimura, K., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .2. NMR studies, *Biochemistry 30*, 11313-11320.

[135] Chen, B. L., and Schellman, J. A. (1989) Low-temperature unfolding of a mutant of phage-T4 lysozyme .1. Equilibrium studies, *Biochemistry 28*, 685-691.

[136] Chen, B. L., Baase, W. A., and Schellman, J. A. (1989) Low-temperature unfolding of a mutant of phage-T4 lysozyme .2. Kinetic investigations, *Biochemistry 28*, 691-699.

[137] Griko, Y. V., Privalov, P. L., Sturtevant, J. M., and Venyaminov, S. Y. (1988) Cold denaturation of staphylococcal nuclease, *Proc Natl Acad Sci USA 85*, 3343-3347.

[138] Pace, N. C., and Tanford, C. (1968) Thermodynamics of unfolding of β-lactoglobulin A in aqueous urea solutions between 5 and 55 degrees, *Biochemistry 7*, 198-206.

[139] Brandts, J. F. (1964) Thermodynamics of protein denaturation .2. Model of reversible denaturation and interpretations regarding stability of chymotrypsinogen, *J Am Chem Soc 86*, 4302-4302.

[140] Tian, J., and Garcia, A. E. (2011) Simulations of the confinement of ubiquitin in self-assembled reverse micelles, *J Chem Phys 134*, 225101.

[141] 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.

[142] 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.

[143] Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyama, S., and Akasaka, K. (2006) Cold denaturation of ubiquitin at high pressure, *Magn Reson Chem 44 Spec No*, S108-113.

[144] 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.

[145] 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.

[146] Adrover, M., Martorell, G., Martin, S. R., Urosev, D., Konarev, P. V., Svergun, D. I., Daura, X., Temussi, P., and Pastore, A. (2012) The role of hydration in protein stability: Comparison of the cold and heat unfolded states of Yfh1, *J Mol Biol 417*, 413-424.

[147] Adrover, M., Esposito, V., Martorell, G., Pastore, A., and Temussi, P. A. (2010) Understanding cold denaturation: The case study of Yfh1, *J Am Chem Soc 132*, 16240-16246.

[148] Martin, S. R., Esposito, V., De Los Rios, P., 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.

[149] 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.

[150] 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 bi-lobed RNA-binding protein, *EMBO J 13*, 205-212.

[151] Sato, S., Luisi, D. L., and Raleigh, D. P. (2000) pH jump studies of the folding of the multidomain ribosomal protein L9: The structural organization of the N-terminal domain does not affect the anomalously slow folding of the C-terminal domain, *Biochemistry 39*, 4955-4962.

[152] 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.

[153] Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution, *Science 310*, 827-834.

[154] 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.

[155] Brimacombe, R., Gornicki, P., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, 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.

[156] Walleczek, J., Martin, T., Redl, B., Stoffler-Meilicke, M., and Stoffler, G. (1989) Comparative cross-linking study on the 50S ribosomal subunit from *Escherichia coli*, *Biochemistry 28*, 4099-4105.

[157] Redl, B., Walleczek, J., Stoffler-Meilicke, M., and Stoffler, G. (1989) Immunoblotting analysis of protein-protein crosslinks within the 50S ribosomal subunit of *Escherichia coli*. A study using dimethylsuberimidate as crosslinking reagent, *Eur J Biochem 181*, 351-356.

[158] Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318*, 571-582.

[159] 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.

[160] Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9, *J Mol Biol 349*, 839-846.

[161] 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.

[162] 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.

[163] Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure, *J Am Chem Soc 132*, 4669-4677.

Chapter 2

Denatured State Ensembles with the Same Radius of Gyration Can Form Significantly Different Long-Range Contacts

Abstract

Defining the structural, dynamic, and energetic properties of the unfolded state of proteins is critical for an in-depth understanding of protein folding, protein thermodynamics, and protein aggregation. This chapter presents an analysis of long-range contacts and compactness in two apparently fully unfolded ensembles of the same protein: the acid unfolded state of the C-terminal domain of the ribosomal protein L9 (CTL9) in the absence of high concentrations of urea as well as the urea unfolded state at low pH. SAXS reveals that the two states are expanded with values of R_g differing by less than 7 %. Paramagnetic relaxation enhancement (PRE) NMR studies, however, reveal that the acid unfolded state samples conformations that facilitate contacts between residues that are distant in sequence while the urea unfolded state ensemble does not. The experimental PRE profiles for the acid unfolded state differ significantly from these predicted using an excluded volume (EV) limit ensemble, but these long-range contacts are largely eliminated by the addition of 8 M urea. The work shows that expanded unfolded states can sample very different distributions of long-range contacts yet still have similar radii of gyration. The implications for protein folding and for the characterization of unfolded states are discussed.

Note: The studies presented in this chapter have been published (Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2014) Denatured State Ensembles with the Same Radius of Gyration Can Form Significantly Different Long-Range Contacts. Biochemisty, 53: 39-47). This chapter contains direct excerpts from the paper with suggestions and revisions by Prof. Daniel P. Raleigh. The generation of the EV model was performed by Dr. Nicholas Lyle and Prof. Rohit V. Pappu. I thank Dr. Bing Shan and Dr. Wenli Meng for helpful discussions and Dr. Marc Allaire for the help of SAXS experiments.

2.1. Introduction

Quantitative characterization of denatured state ensembles (DSE) of proteins, also referred as the protein unfolded or denatured state, is important for understanding the mechanism of protein folding. The DSE is the starting point of protein folding, the thermodynamic reference state for protein stability, and it can be targeted by rational protein design.[1-7] Studies of DSEs can also reveal factors that impact protein misfolding, and modulate the tendency for protein aggregation *in vitro* and *in vivo* and amyloid formation.[8-12] The exploration of the mechanisms and biological function of intrinsically disordered proteins (IDPs) largely depends on the characterization of the properties of unstructured and partially structured states and therefore has much in common with studies of the DSE.[13, 14]

The properties of the DSE can vary considerably depending upon the conditions used to populate it. Under near native conditions the DSE can be compact with significant residual structure, while more expanded and less structured DSEs are usually populated under strongly denatured conditions. Small angle X-ray scattering (SAXS) is frequently used to study the overall compactness of the DSEs and provides the radius of gyration (R_g) and in favorable cases more information.[15-20] DSEs which have the same value of R_g are often assumed to be similarly unfolded.[16, 17, 21]

In strongly denaturing conditions, the DSE expands to make favorable interactions with the solvent, and the R_g of proteins without disulfide cross links follows a non-trivial

power law relationship, which scales with the number of amino acids in the peptide chain, N, as $N^{0.59}$.[16, 17, 22] Similar scaling behavior is observed for polymers modeled as self-avoiding random walks.[23] Observation of an R_g value consistent with this scaling is often taken to mean a protein is fully unfolded, however, this scaling does not preclude the possibility of detectable, low-likelihood native and non-native contacts within expanded DSEs, even under strongly denaturing conditions.[2-6, 13, 24-37] However, it is unclear if different DSEs generated for the same protein under different conditions, all with similar R_g values will exhibit similar patterns of, and likelihoods for native and non-native contacts. This issue is important because such contacts might contribute directly to folding, and might influence the tendency to aggregate. It is also important because it potentially highlights the need to go beyond measurements of R_g alone as an adjudicator of the degree of unfoldness and as a descriptor of unfolded states.

Here we examine the 92-residue C-terminal domain of the ribosomal protein L9 in the acid-induced DSE and in the low-pH urea-induced DSE, in order to determine if the conformational ensembles are different between different DSEs that have similar R_g values. The domain, denoted as CTL9, has a mixed parallel, anti-parallel β-sheet, two loops with partial α-helix 3_{10} turns, and two α-helices (**Figure 2-1**). CTL9 can be denatured by lowering the pH, or by adding denaturant, and folds in a two state fashion.[38-42] Previous studies have used NMR chemical shift analysis to show that the two ensembles differ in their residual secondary structure, but have similar values of the hydrodynamic radius, R_h.[41, 42] In particular, more helical structure was detected in the acid unfolded state.[41] R_h is a useful parameter for probing compactness, but it is

determined indirectly from NMR diffusion measurements and involves comparison to an added internal standard and requires assumptions about the hydrodynamic properties of the ensemble. R_g measurements offer an alternative probe of the compactness of the DSE, but involve different assumptions than the diffusion measurements and are, in some sense, a more direct measure of the ensemble. The previous NMR studies and hydrodynamic measurements were unable to define the nature of any long-range contacts. This information is required for a full description of the DSE and one might expect that states which are equally expanded should have a similar level of long-range contacts. The work presented in this chapter shows that this is not the case, and also directly demonstrates that states which have very similar values of R_g, which furthermore are consistent with the scaling expected for expanded highly unfolded states, can differ dramatically in the extent of detectable long-range contacts.

NMR paramagnetic relaxation enhancement (PRE) measurements were used to detect long-range contacts and R_g was measured using SAXS for two different DSEs of CTL9. The two ensembles have very similar R_g distributions, but exhibit significantly different patterns of long-range contacts. The urea unfolded DSE contains few long-range contacts and is well described by an excluded volume model while the acid unfolded DSE displays extensive long-range interactions.

2.2. Materials and methods

2.2.1. Protein expression and purification

^{15}N-labeled wild type CTL9 and the cysteine mutants were over expressed in the *Escherichia coli* strain BL21 cells M9 minimal media and purified as described previously.[38] 0.8 g/L $^{15}NH_4Cl$ was used as the sole source of the nitrogen in the expression of ^{15}N-labeled proteins. The cells were grown at 37 °C until the optical density at 600 nm reached 0.8-0.9, followed by 1 mM IPTG (isopropylthio-β-D-galactoside) induction for 4 hours at 37 °C. Cells were harvest and lysed by sonication, and the lysate was centrifuged for 1 hr at 10,000 g to remove the cell debris. The supernatant was loaded onto a SP-Sepharose fast flow ion-exchange column (GE Healthcare). 20 mM Tris buffer at pH 7.4 was used equilibrium buffer, and the protein was eluted with a 0 to 2 M NaCl gradient. Proteins were further purified with the reverse phase HPLC using a C8 column (Vydac). An A-B gradient was created using buffer A containing 99.9% H_2O and 0.1% TFA (trifluoroacetic acid) and buffer B containing 90% acetonitrile, 9.9% H_2O, and 0.1% TFA. The protocol used is 100% A for 10 mins after injecting the sample followed by 25-55% B in 60 mins. CTL9 eluted around 40-42% B. The typical yield for CTL9 Cys mutants are ~ 20-25 mg/L. The identity of the protein was confirmed by MALDI-TOF mass spectroscopy.

2.2.2. Small angle X-ray scattering (SAXS) experiments and data analysis

Experiments were performed on the acid-induced DSE and on the urea-induced DSE. Samples of CTL9 wild type were prepared using the same buffers used for the NMR PRE experiments. Scattering experiments were performed at beamline X9 at Brookhaven National Laboratory, National Synchrotron Light Source I (Upton, New York, USA).

Samples were injected continuously into a 1 mm capillary during the measurement at a rate of 0.67 μL/s in order to avoid radiation damage. The concentrations of the acid-induced and the urea-induced CTL9 DSE samples were at 2.5 mg/mL. The exposure time for each measurement was 30 s. Each sample was measured three times and then averaged before data analysis. The program pyXS (http://www.bnl.gov/ps/x9/software/pyXS.asp) was used for buffer subtraction, the data was then processed using standard procedures by the program PRIMUS.[43] The scattering patterns were analyzed using the ensemble optimization method (EOM) to determine the size distribution for the DSE in urea and in acid. EOM creates a pool of structures based on the protein primary sequence, and then an ensemble of protein structures that best fits the experimental data is selected.[44]

2.2.3. Circular dichroism (CD) spectroscopy

Thermal unfolding experiments of wild type CTL9 and the six cysteine mutants were performed with an Applied Photophysics Chirascan CD instrument, over the temperature range of 4 to 92 °C with an interval of 2 °C. The absorbance of the spin label prevents CD monitored unfolding studies of the derivatized mutants. The protein concentration was 16 to 19 μM, in 10 mM MOPS (3-(N-morpholino) propanesulfonic acid) and 150 mM NaCl at pH 7.5. The signal at 222 nm was monitored. All thermal unfolding curves were analyzed by nonlinear least-squares fitting using SigmaPlot 10.0.

A pH titration curve was measured for wild type CTL9 at 25°C, in 10 mM sodium acetate and 150 mM NaCl, over the range of pH 1.8 to 12.0. Urea unfolding data were collected at 25°C on wild type CTL9, using an AVIV Instruments model 202SF CD instrument. The native protein sample was titrated with urea in 0.25 M increments until the final urea concentration reached 10 M. The buffer was the same as used for the thermal unfolding experiments, and the urea concentration was determined by measuring the refractive index.

2.2.4. Paramagnetic relaxation enhancement (PRE) experiments

Single cysteine mutants of CTL9 were dissolved in 600 uL of NMR buffer, containing 10% D_2O. 5 uL of the 300 mM MTSL ((1-oxyl-2,2,5,5-teramethyl-3-pyroline-3-methyl)methanesulfonate) stock solution was added to the NMR sample and then the pH was adjusted to 2.0. After 12 hrs incubation at room temperature, the sample was loaded onto a Sephadex G25 column in order to remove the excess MTSL. The completeness of labeling was assessed using HPLC, and by the intensity of the Cys NMR peaks. The Cys resonance is bleached in the labeled sample due to the attachment of MTSL, but is not bleached in the unlabeled sample. Diamagnetic samples were prepared by adding TCEP (tris(2-carboxyethyl)phosphine) to a final concentration of 2.5 mM. Urea-denatured samples were prepared by adding urea to the samples immediately after the desalting process. The final urea concentration was 8 M, determined by measuring the refractive index. The final protein concentration was 250 uM.

^{15}N-^{1}H correlated heteronuclear single coherence (HSQC) experiments were performed on both samples in paramagnetic and diamagnetic forms. The spectra were collected at 25 °C, using 1024 × 256 complex points with 8 scans per increment. The spectral widths were 6009.6 Hz and 1517.8 Hz for the ^{1}H and the ^{15}N dimensions respectively. 3D TOCSY-HSQC spectra were collected on the ^{15}N-labeled diamagnetic samples to confirm peak assignments. The mixing time was 75 ms, and the data matrix was 1024 (direct ^{1}H dimension) × 128 (indirect ^{1}H dimension) × 128 (^{15}N dimension). Spectra widths for the direct ^{1}H dimension, the indirect ^{1}H dimension and the ^{15}N dimension were 6009.6 Hz, 6009.6 Hz and 1500.0 Hz, respectively. Tables of NMR assignments of the six Cys mutants of CTL9 are included in the appendix of this thesis.

2.2.5. NMR data processing and determination of PREs

All spectra were processed using NMRPipe and visualized via NMRView. [45, 46] The ratio of intensity for a particular residue was calculated as I_{para}/I_{dia}, where I_{para} is the intensity of the paramagnetic form, with spin label MTSL; and I_{dia} was the peak intensity for the diamagnetic sample, with the spin label MTSL reduced. Lower intensity ratios correspond to a larger PRE effects and indicate an interaction with the spin label. Tables of the PRE ratios from the six Cys mutants of CTL9 are included in the appendix of this thesis.

Two different models were used to calculate baseline PRE effects for a highly unfolded chain: a Gaussian distribution model and an excluded volume (EV) model. The EV calculations were performed by the Pappu Laboratory.

2.2.6. Generation of the Gaussian distribution model

In this model, a Gaussian distribution of root-mean-square end-to-end distances is used to describe the distances between each residue and the site of spin label:[34, 36]

$$\langle r^2 \rangle = nl^2 \left[\frac{1+\alpha}{1-\alpha} - \frac{2\alpha(1-\alpha^n)}{n(1-\alpha)^2}\right] \quad \text{(eq 2-1)}$$

n is the number of residues between residue i and the site of spin label, r is the end-to-end distance between a residue and the site of the spin label, l (3.8 Å) is the link length of the chain, 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 the statistical segment lengths in poly-L-alanine. R_{2P} is the contribution of the paramagnetic relaxation enhancement to the transverse relaxation rate, and was calculated using:

$$R_{2P} = \frac{K}{r^6}\left(4\tau_c + \frac{3\tau_c}{1+\omega_H^2\tau_c^2}\right) \quad \text{(eq 2-2)}$$

Here r is the distance between a given residue and the site of spin-label, K is a constant equal to 1.23×10^{-32} cm^6 s^{-2}, τ_c is the effective correlation time, taken here as 3.8 ns, and ω_H is the proton Larmor frequency. The value for the effective correlation time was

chosen based on NMR relaxation measurements. The peak intensity ratios between the paramagnetic and diamagnetic forms were calculated using:

$$\frac{I_P}{I_D} = \frac{R_{2D}\exp(-R_{2P}t)}{R_{2D}+R_{2P}} \quad \text{(eq 2-3)}$$

R_{2D} is the transverse relaxation rate of the backbone amide protons in the diamagnetic form. The average value was determined to be 13.5 s^{-1} using a 1D NMR, t is set to be 12 ms, equal to the total duration of the INEPT delays in the HSQC pulse sequence.

2.2.7. Genration of the EV ensemble

Calculations were performed using the CAMPARI software package for Metropolis Monte Carlo (MC) simulations based on the ABSINTH implicit solvation model and underlying forcefield paradigm,[47] and parameters from the OPLSS-AA/L molecular mechanics forcefield, specifically parameters in abs3.2_opls.prm.[48] The internal degrees of freedom included the backbone φ, ψ, ω and sidechain χ dihedral angles. More details regarding the MC moveset are detailed in Meng, et. al.[49] Results were averaged over ten independent simulations. Each simulation underwent 4×10^7 MC moves, not including an equilibration phase of 1×10^5 moves. The radius of gyration over CTL9 was accumulated every 1×10^4 MC moves.

2.3. Results

The folding of CTL9 has been characterized previously in terms of its kinetics and thermodynamics, and the cold denaturation of a destabilized mutant has been probed.[38, 40, 50-53] The stability and folding rate of CTL9 are strongly dependent on pH, due in part to the three His residues in the protein and the domain can be unfolded by lowering the pH, as well as by adding denaturant. CD monitored titration curves indicate the transformation from the native folded state to the DSE is well fit by a two-state model (**Figure 2-2**). The acid unfolding transition is completed by pH 2.8 in the absence of urea and the urea unfolding transition is complete by 5.5 M urea at pH 5.6 as judged by CD. The degree of secondary structure in the two DSEs is difficult to determine from CD since the strong absorbance of urea limits the wavelength range that is accessible. In addition, the CD signal from short α-helices can differ from standard spectra.[54] However, previous studies used NMR to probe residual secondary structure in both of these states. $C_\alpha{}^1H$, $^{13}C_\alpha$, and $^{13}C_\beta$ chemical shifts were analyzed and revealed that in the acid unfolded state there is a modest, but non-zero, propensity to preferably populate the helical region of the Ramachandran plot for those residues which are helical in the native state. In contrast, this propensity to form secondary structure was significantly reduced in the low pH urea unfolded state.[41]

2.3.1. SAXS experiments show that the urea-induced and acid-induced DSEs are expanded

The R_g measured for the native state of CTL9 is 14.5 ± 0.3 Å.[52] The values of R_g for the acid and urea DSEs, determined from the Guinier plot are 30.8 ± 1.6 Å and 32.9 ± 1.5 Å,

respectively. The difference between the two R_g values is 6% and is statistically insignificant given the intrinsic uncertainty associated with each value. The value of R_g predicted for a fully unfolded 92-residue polypeptide based on empirical scaling relationships is 28.9 ± 4.6 Å.[16] By this criteria both ensembles are classified as expanded unfolded states. We also fit the scattering patterns using the Ensemble Optimization Method (EOM).[44] The EOM algorithm builds a pool of structures based on the primary sequence of the target protein, and a series of theoretical scattering curves are generated. A combination of the generated structures is used to generate an ensemble of representative structures that reproduce the experimental data. The average R_g of the DSE can be deduced from the EOM fitting of the experimental data. The average R_g values estimated using this method are within 8% of each other, 30.8 Å and 33.5 Å for the acid and urea-induced DSE, respectively (**Figure 2-3**), and the widths of the distribution are 8.0 and 9.0 Å.

2.3.2. PRE studies reveal long-range contacts in the acid-induced DSE, but not in the urea-induced DSE

We used paramagnetic relaxation enhancement (PRE) experiments in order to obtain information about long-range contacts. PRE is frequently applied to probe long-range interactions in the DSE and is sensitive to distances up to 20 Å.[34, 36, 49, 55-70] Spin labels were attached to six sites using single cysteine mutants. The sites are all surface exposed in the native state (**Figure 2-1**): E61 near the N-terminus of the first α-helix, K74 at the C-terminus of the same α-helix, K96 near the N-terminus of the second α-helix, K109

and D119 in the loop region between the second α-helix and the second β-strand, together with K149 which lies at C-terminus. The mutants are denoted as E61C-CTL9, K74C-CTL9, K96C-CTL9, K109C-CTL9, D119C-CTL9, and K149C-CTL9, respectively. The CD spectra of the mutants are similar to that of wild type CTL9 (**Figure 2-4**). Comparison of the thermodynamic properties determined from the thermal unfolding show that the cysteine mutants have similar values of T_m and ΔH^o with respect to wild type CTL9 (**Table 2-1** and **Figure 2-5**), although the T_m of D119C-CTL9 is somewhat depressed compared to wild type. The peaks in the ^1H-^{15}N HSQC of the acid- and urea-induced DSE of the mutants do not shift their positions, relative to wild type, aside from the mutated residue (**Figure 2-6** – **Figure 2-12**). Tables of NMR assignments of the six Cys mutants of CTL9 are included in the appendix of this thesis. Thus all of the available data suggest that the mutations do not alter the properties of the native state or the DSE.

^1H-^{15}N HSQC spectra were collected for the six mutants with spin labels (paramagnetic state) and without spin labels (diamagnetic state), in 8 M urea. The cross peak intensity ratios I_{para}/I_{dia} were calculated and plotted against the corresponding residue number (**Figure 2-13**). A table of the PRE ratios from the six Cys mutants of CTL9 is included in the appendix of this thesis. The expected PRE profile for a random coil model was generated using two different models. We used a Gaussian model, in which there is a Gaussian distribution of distances between the spin label site and each residue of the protein,[34, 36] and an excluded volume (EV) model. The Gaussian chain model has been widely used to benchmark PRE studies of unfolded states owing to its simplicity. However the model lacks any detail. Quantitative descriptions of chain statistics for

polymers in good solvents rely on the so-called EV limit as a reference state and this is true for denatured state ensembles as well.[71-77] The EV model is generated using an all atom representation of the chain with only excluded volume interactions, which also takes into account the size and flexibility of the side chain linked spin label, the EV ensemble corresponding to the wild-type sequence was used.[73] EV ensembles were generated for CTL9 using atomistic descriptions of proteins and all non-bonded interactions excepting steric repulsions were ignored.

The observed PRE effects for all the mutants in 8 M urea correlate very well with the random coil models, especially the EV model, indicating that the urea-induced DSE of CTL9 in 8 M urea at pH 2.5 behaves like a highly unfolded chain.

The same strategy was used to study the acid-induced DSE. However, in this case, clear deviations from both models were observed for four of the six spin-labeled mutants (**Figure 2-13**). The two exceptions are the labels near the N- and C- termini. The other four PRE profiles display significant differences between the experimental and calculated profiles (**Figure 2-14**). Significant PRE effects, arbitrarily defined here as a value of 0.5 or less for the ratio I_{para}/I_{dia}, are detected for sites as far as 30 residues away from the spin label. Even longer-range effects are observed for some of the spin labels, including residues 74 and 109.

2.4. Discussion

The acid- and urea-induced DSE of CTL9 differ significantly in the pattern of long-range contacts. The differences in the intensity ratios between the PRE profiles for the two DSE's are compared in Figure 4 as difference plots, positive values indicate stronger PRE effects in the acid-induced DSE of CTL9. For the N- and C-terminus, residues 61 and 149, the difference in the intensity ratios are close to zero, indicating the there are no obvious changes in the acid- and urea-induced PRE profile. Positions 74, 96, 109, and 119 display positive values, and clearly illustrate the deviations between the two DSEs. The analysis of the acid-induced DSE of CTL9 clearly confirms our previous results with NTL9 that showed long-range contacts can form in a highly expanded DSE.[49]

There are also differences in the local secondary structure propensities of the urea and acid unfolded DSEs of CTL9 despite their similar values of R_g. NMR studies have shown that the urea-induced DSE of CTL9 contains very little residual secondary structure, while in the acid unfolded DSE there are two regions which have a modest propensity to preferentially populate helical angles which are helical in the native state. The average SSP scores for those regions haven been reported to be 0.28 ppm for the α-helix-1 and -0.32 for the α-helix-2.[41] The SSP score is an empirical NMR parameter that is used to estimate secondary structure. It is derives from a combination of secondary shifts, typically $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$ chemical shifts. A value of 1 indicates fully formed helix, and a value of -1 indicates fully formed β-sheet.[78] However, there is no obvious significant correlation between the regions of the polypeptide chain that exhibit an increased propensity for helical angles and strong PRE effects. The present study, together with the previous work, shows that DSEs with similar R_g values can differ

significantly in their patterns of long-range contacts and as well as their propensities to form local structure.

The data obtained on CTL9 further demonstrates that $R_{g,}$ while a very useful measure of unfolded state dimensions, should not be used as the sole criterion to judge if a protein is "fully unfolded", or if longer-range contacts are absent. This may be relevant for SAXS and FRET studies of DSE's. There are examples where SAXS indicates a highly unfolded state, but FRET reveals apparent long-range contacts.[79] The present analysis, together with other studies suggests that these observations can be comparable.[49, 74, 75, 80]

The data collected on the urea-induced DSE of CTL9 shows that it lacks detectable long-range contacts, but this should not be interpreted to mean that all expanded states which lack secondary structure are devoid of long-range contacts. The N-terminal domain of L9 provides a counter example.[49] The urea induced DSE of that domain transiently populates long and medium range contacts, but the distribution of internal distances is still consistent with $N^{0.59}$ scaling and the value of R_g is consistent with an expanded DSE. Reduced hen egg white lysozyme offers another example. The urea induced DSE appears to contain transient clusters of hydrophobic residues, as judged by ^{15}N-NMR relaxation measurements, and they can be modulated by mutation.[35] In contrast, ^{15}N-NMR relaxation measurements on the urea induced DSE of other proteins suggests that these sorts of contacts can be less populated in other proteins.[81] Collectively, the available data in the literature argues that the formation of long-range contacts in expanded unfolded states depends on the protein primary sequence and is not a generic property of all

polypeptides. There is a connection with emerging studies of IDP's, Recent work has revealed that the patterning of residues, i.e. the specific distribution of polar and hydrophobic residues, significantly influences the properties of IDP's and provides a more precise description of their behavior than simple correlations based on mean hydrophobicity and average net charge.[82]

Table 2-1. $\Delta H^o(T_m)$ and T_m for the wild type CTL9 and single cysteine mutants, measured at pH 7.5, in 10 mM MOPS and 150 mM NaCl.

Protein	T_m **(°C)**	**ΔH°(kcal/mol)**
Wild type CTL9	78.4 ± 0.4	67.3 ± 2.3
E61C-CTL9	76.2 ± 0.4	60.0 ± 1.9
K74C-CTL9	75.4 ± 0.2	65.1 ± 1.7
K96C-CTL9	78.2 ± 0.3	61.5 ± 1.7
K109C-CTL9	78.2 ± 0.3	65.8 ± 1.8
D119C-CTL9	69.9 ± 0.2	69.5 ± 2.2
K149C-CTL9	77.2 ± 0.3	67.1 ± 2.2

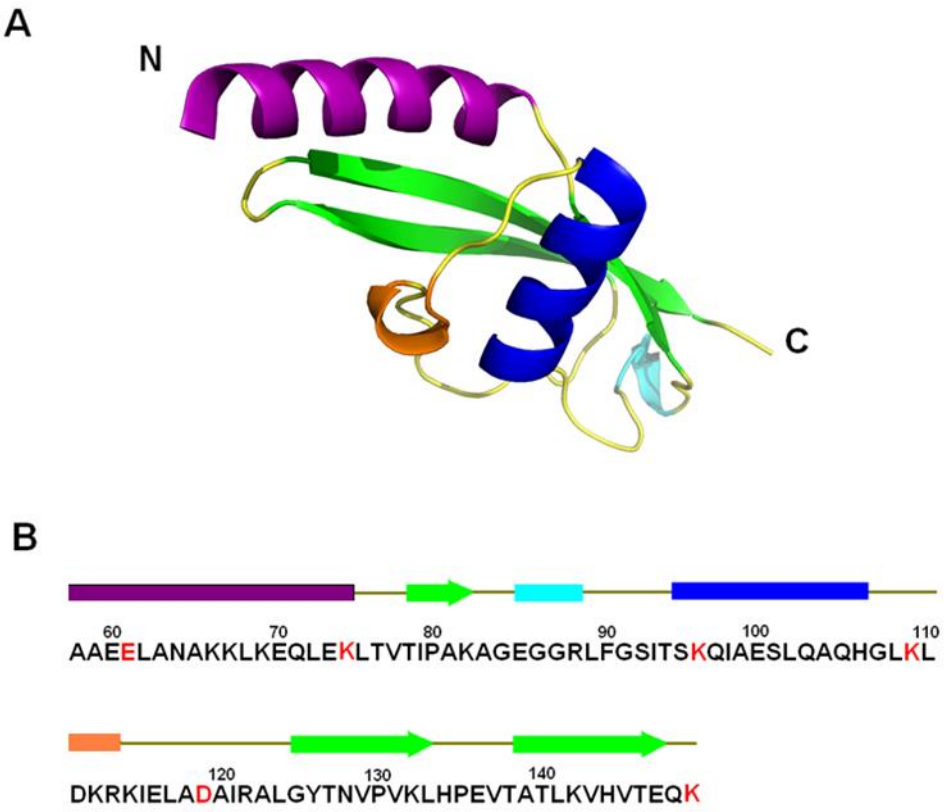


Figure 2-1. (A) The structure of the C-terminal domain of the ribosomal protein L9 (CTL9), corresponding to residues 58 to 149 of L9, shown as a ribbon diagram, PDB data bank entry 1DIV. The N- and C-terminus are labeled. (B) The primary sequence of CTL9. The sites used to attach the spin labels are colored red. A schematic representation of the secondary structure elements of CTL9, using the same color coding as employed for the ribbon diagram is shown above the sequence. Helices are represented as rectangles, β-sheets as arrows.

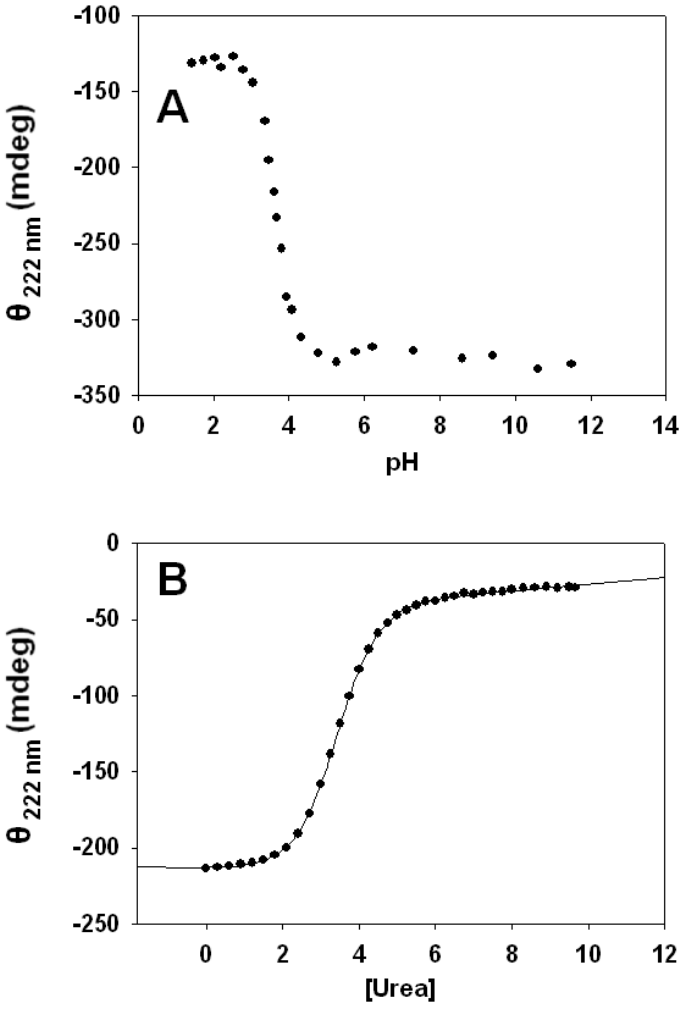


Figure 2-2. pH and urea induced unfolding of CTL9. (A) CD signal as a function of pH, samples contained 10 mM sodium acetate and 150 mM NaCl at 25 ºC. (B) CD signal as a function of urea concentration, samples contained 10 mM MOPS and 150 mM NaCl, pH 5.6, 25 ºC. Both samples were 19 μM CTL9.

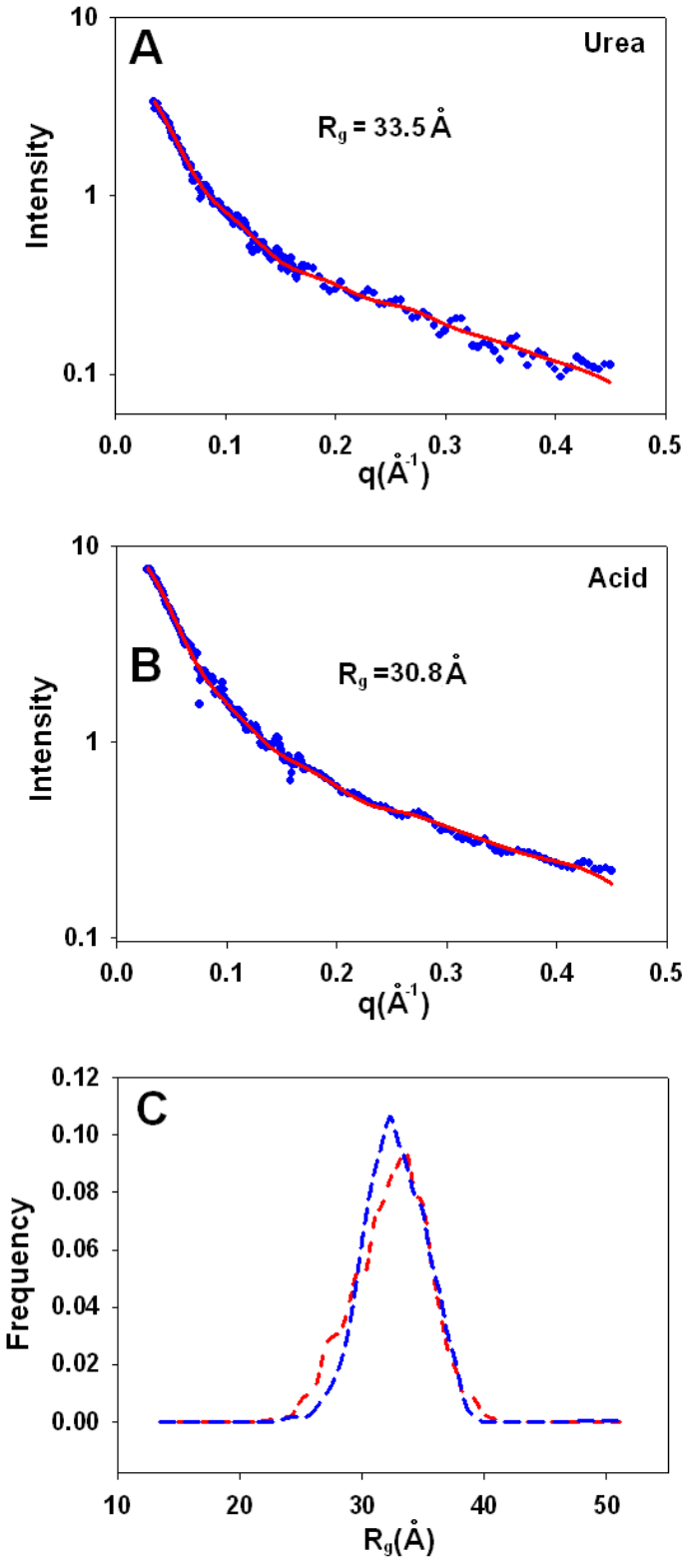

A
Urea
R_g = 33.5 Å
Intensity
10
1
0.1
0.0
0.1
0.2
0.3
0.4
0.5
q(Å^{-1})
B
Acid
R_g =30.8 Å
Intensity
10
1
0.1
0.0
0.1
0.2
0.3
0.4
0.5
q(Å^{-1})
C
Frequency
0.12
0.10
0.08
0.06
0.04
0.02
0.00
10
20
30
40
50
R_g(Å)

Figure 2-3. The acid and urea unfolded states of CTL9 are expanded. (A) SAXS scattering curve from a sample of CTL9 in 8 M urea, at pH 2.5. (B) SAXS scattering curve from a sample of CTL9 at pH 2.0. The experimental data are shown as blue dots and the fit using Ensemble Opitimiztion Method is shown as a red line. (C) The R_g distribution for the DSE in urea (--) and in acid (--), calculated using the Ensemble Opitimiztion Method. The average R_g values are 33.5 Å and 30.8 Å, respectively.

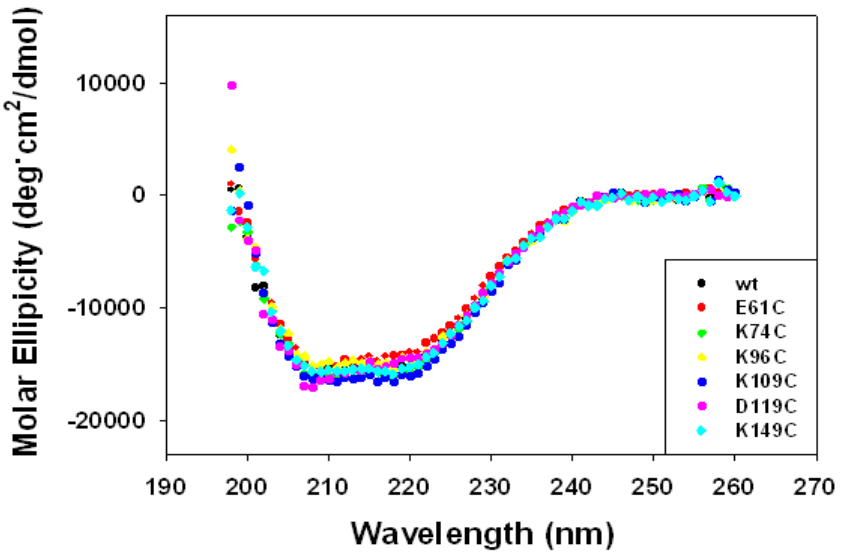


Figure 2-4. CD spectra of wild type CTL9 and the six cysteine mutants. Experiments were conducted at pH 7.5, in 10 mM MOPS and 150 mM NaCl.

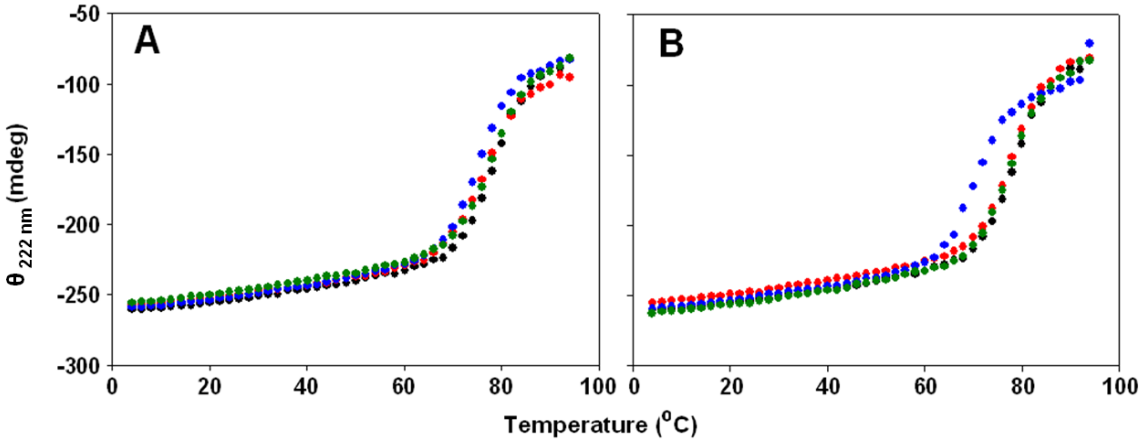


Figure 2-5. Thermal denaturation monitored by CD. (A) E61C (•), K74C (•), and K96C (•) are compared with wild type CTL9 (•). (B) K109C (•), D119C (•), and K149C (•) are compared to wild type CTL9 (•). Experiments were conducted at pH 7.5, in 10 mM MOPS and 150 mM NaCl.

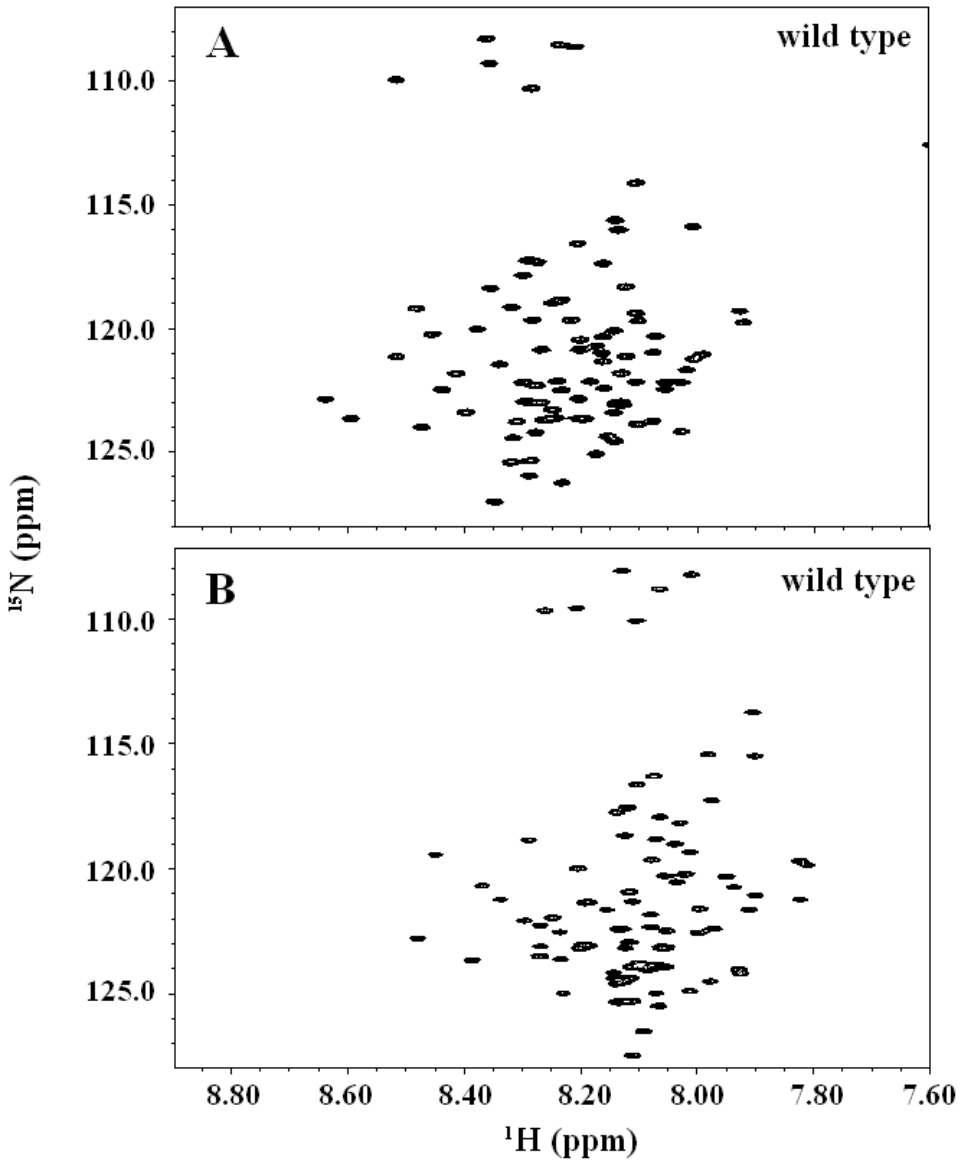


Figure 2-6. (A) ^{1}H-^{15}N HSQC spectra of wild type CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of wild type CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

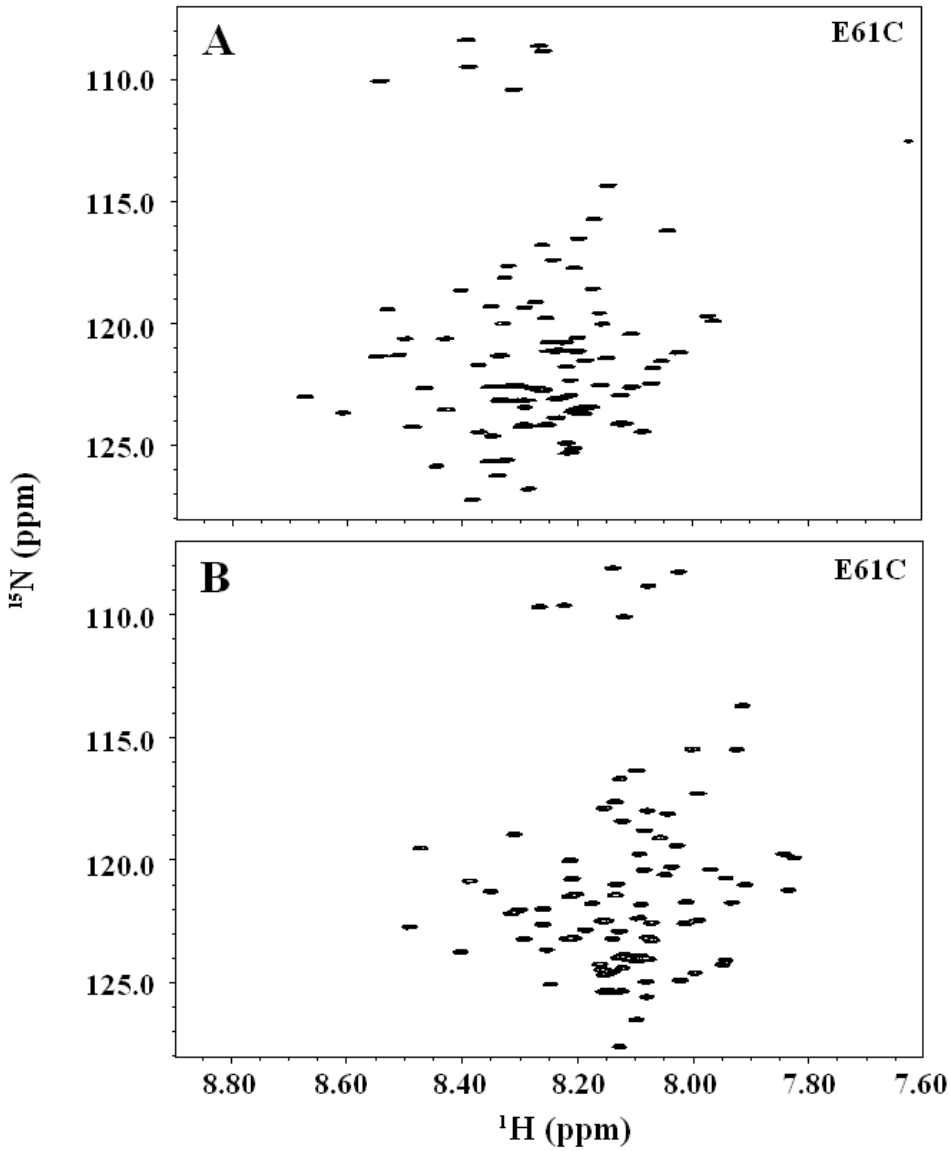


Figure 2-7. (A) ^{1}H-^{15}N HSQC spectra of E61C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of E61C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

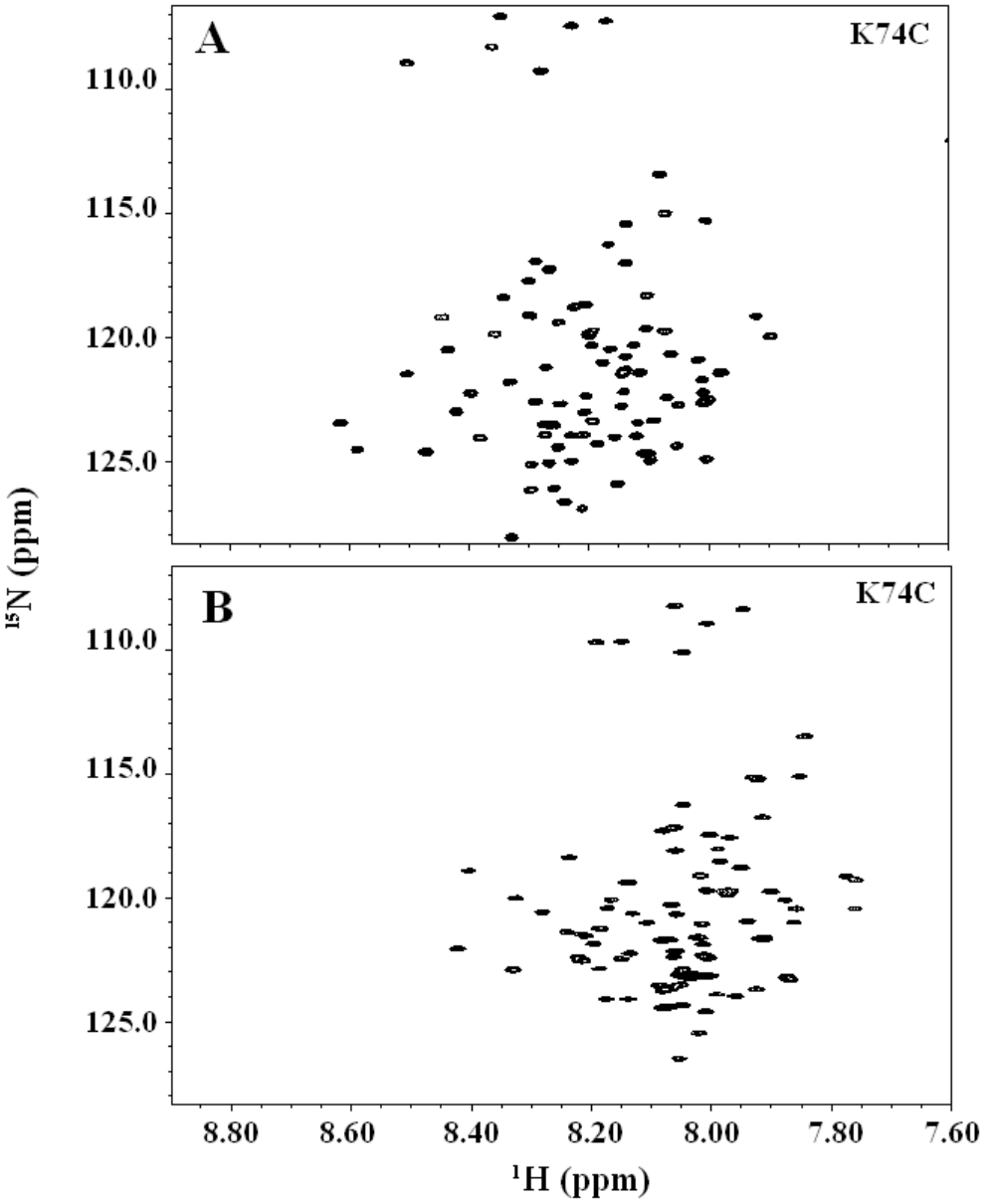


Figure 2-8. (A) ^{1}H-^{15}N HSQC spectra of K74C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of K74C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

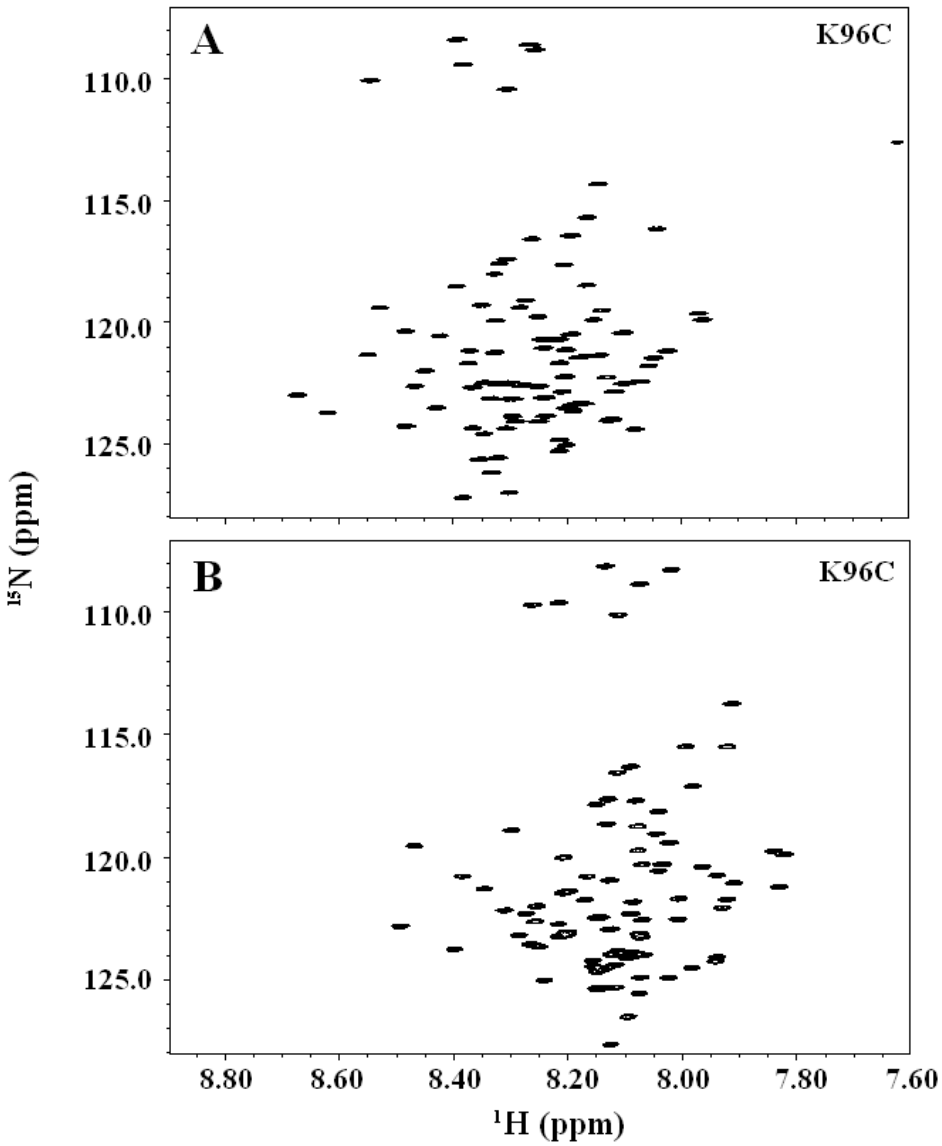


Figure 2-9. (A) ^{1}H-^{15}N HSQC spectra of K96C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of K96C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

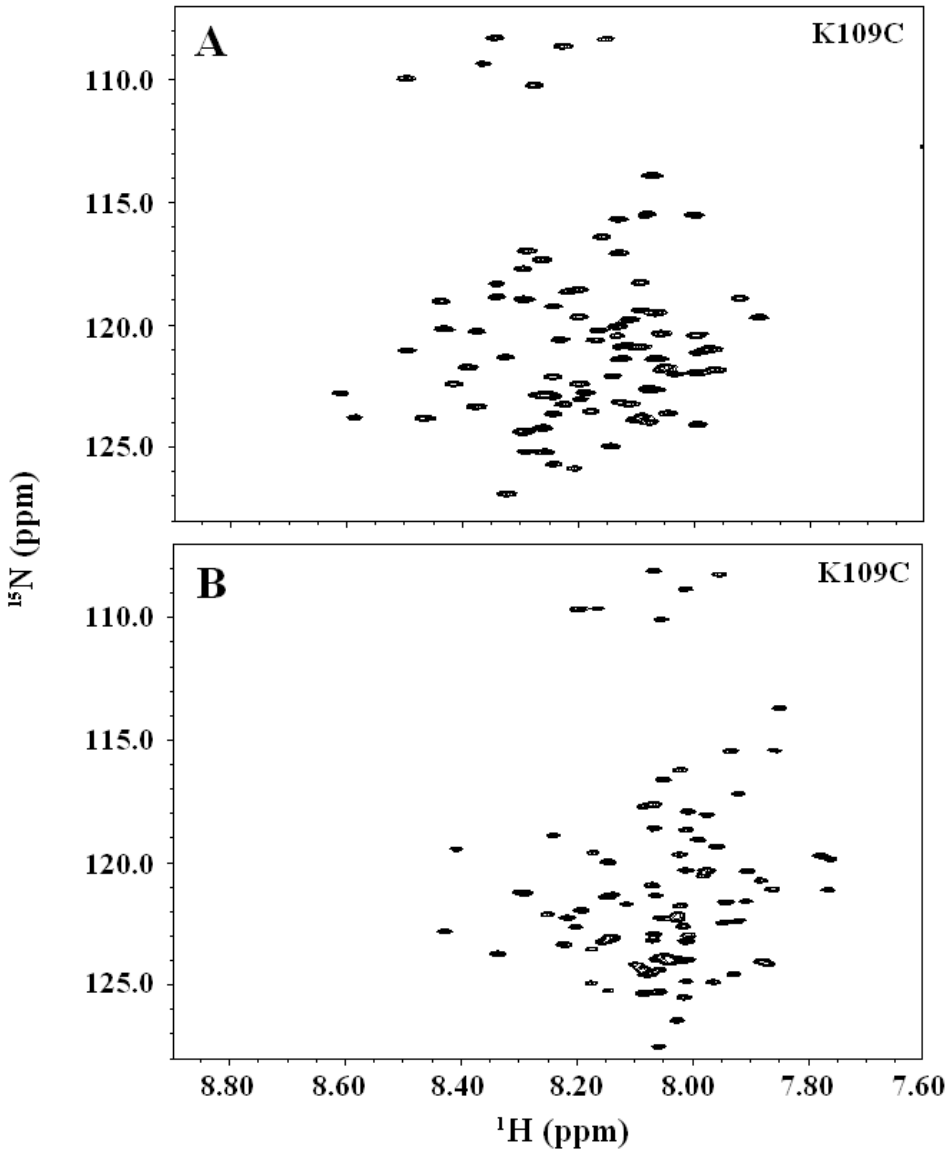


Figure 2-10. (A) ^{1}H-^{15}N HSQC spectra of K109C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of K109C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

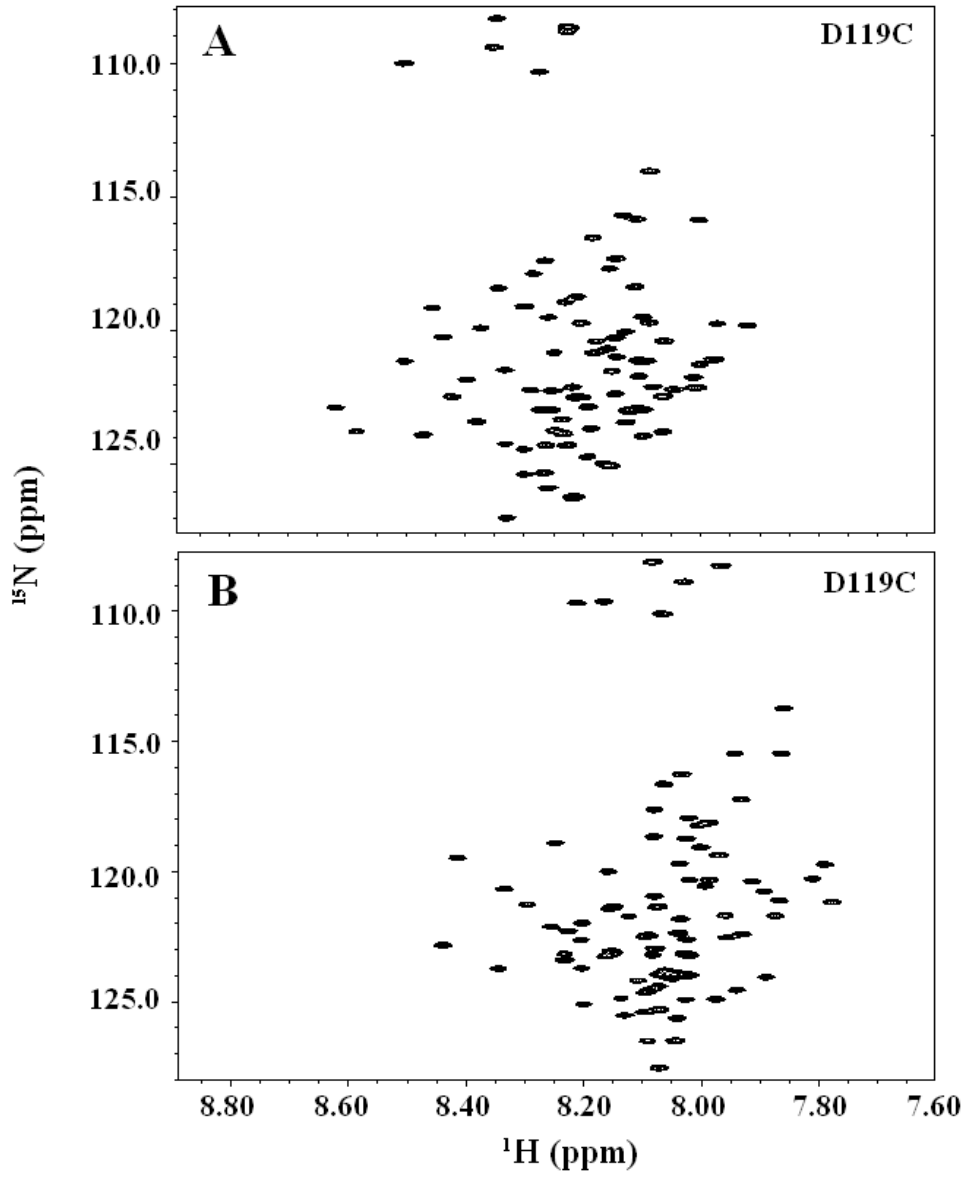


Figure 2-11. (A) ^{1}H-^{15}N HSQC spectra of D119C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of D119C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

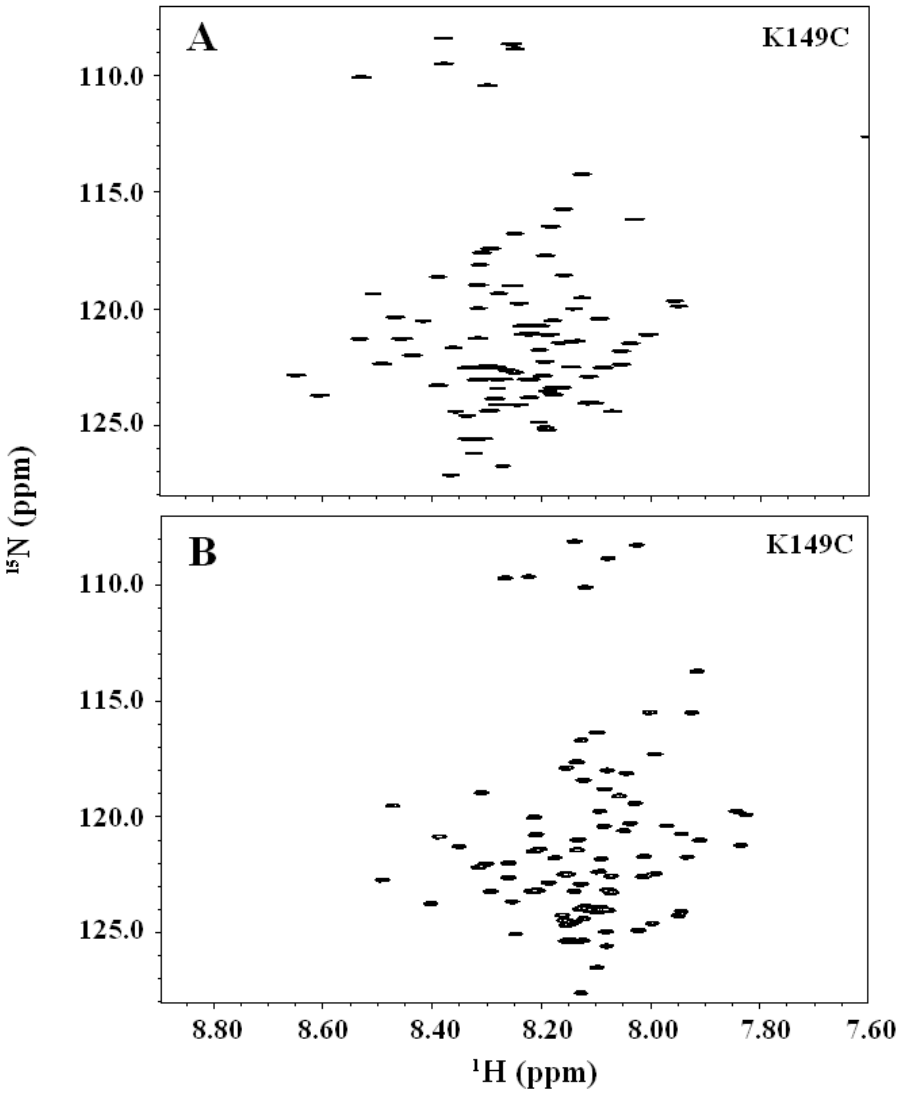


Figure 2-12. (A) ^{1}H-^{15}N HSQC spectra of K149C-CTL9 DSE at pH 2.0 in the absence of urea. Spectra were recorded at 25ºC, with the pH adjusted to 2.0 by adding HCl. (B) ^{1}H-^{15}N HSQC spectra of K149C-CTL9 DSE in urea. Spectra were recorded at 25ºC, in 8 M urea at pH 2.5.

E61C
K74C
K96C
Urea
Acid
Intensity Ratio
Residue Number
0.0
0.2
0.4
0.6
0.8
1.0
1.2
60
80
100
120
140

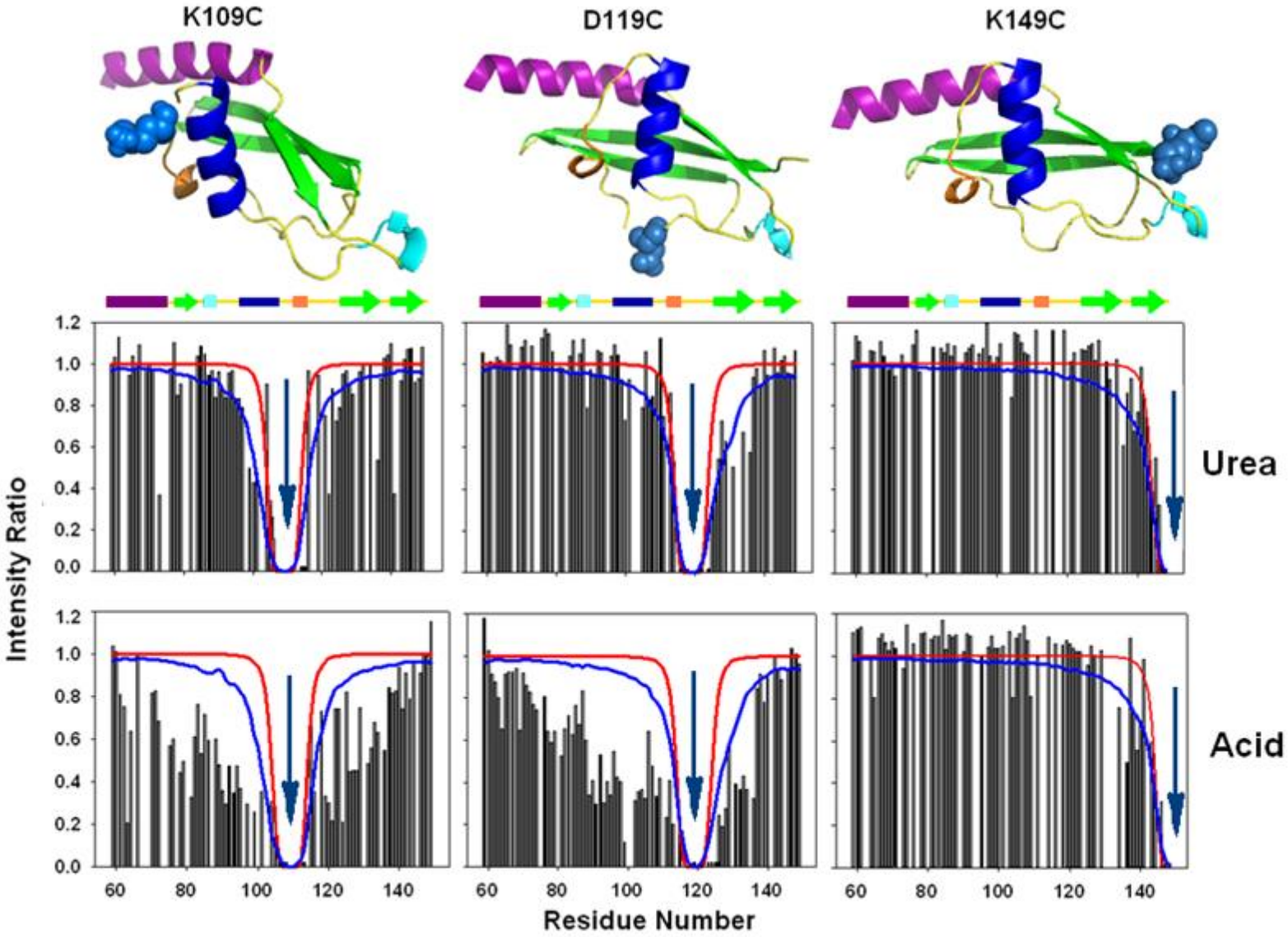

K109C
D119C
K149C
Intensity Ratio
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Urea
Acid
60
80
100
120
140
Residue Number

Figure 2-13. Paramagnetic relaxation enhancement experiments show that there are significant deviations from random coil behavior for the acid unfolded state, but not for the urea unfolded state. The histograms display the intensity ratio of the ^{1}H-^{15}N cross peaks in the HSQC spectra. The dark blue arrow (↓) indicates the site of attachment of the spin label. The solid red curve represents the values predicted by the Gaussian model and the solid blue curve the values predicted by the Excluded Volume model. Ribbon diagrams of CTL9 are shown for each mutant indicating the site of attachment of the spin label. Experiments were carried in 8 M urea at pH 2.5 and at pH 2.0 in the absence of urea.

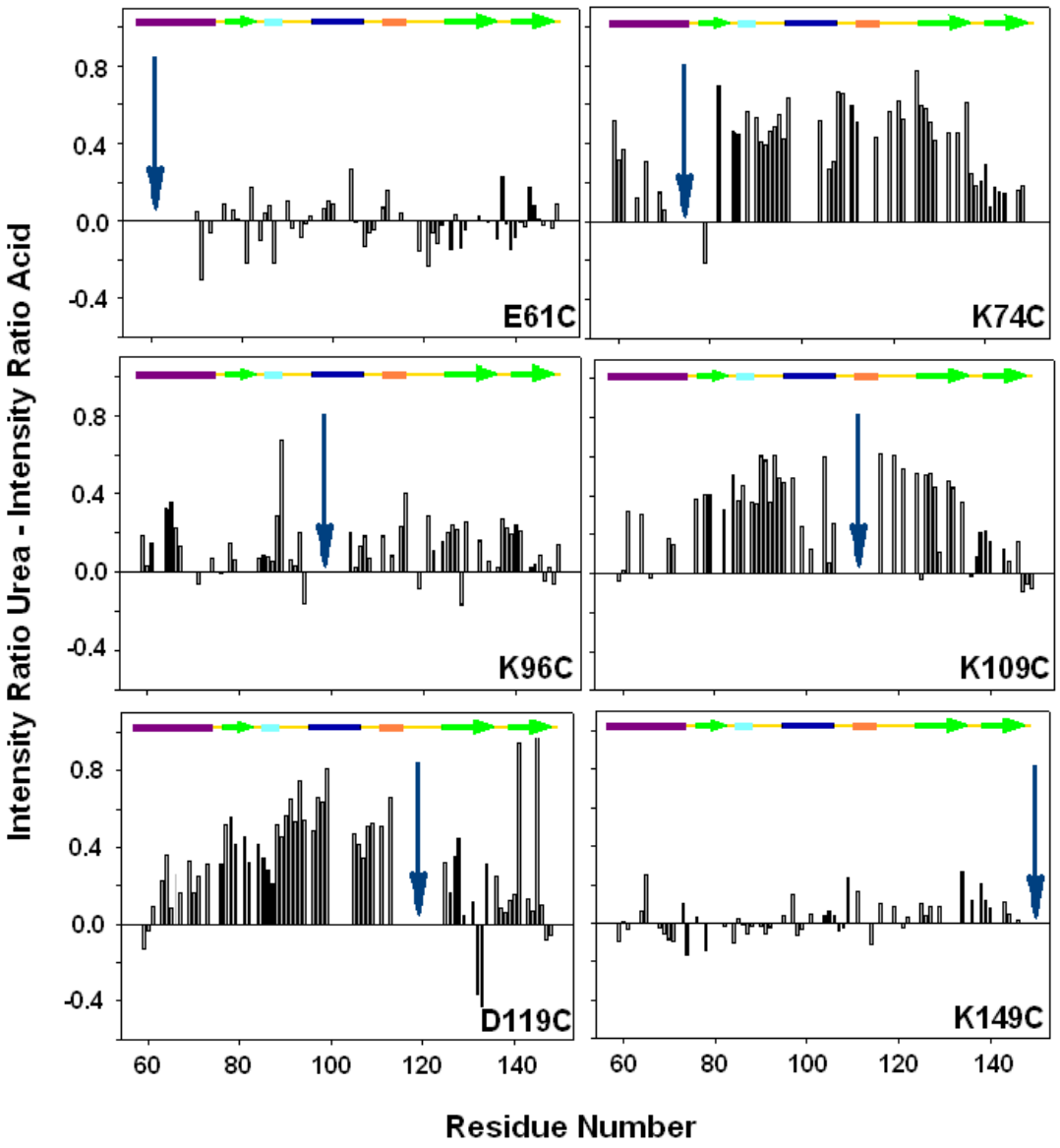


Figure 2-14. Plots of the difference between the PRE effects in the urea DSE and the acid DSE. Data are plotted as the PRE ratio in the urea-induced DSE minus the PRE ratio in the acid-induced DSE. The dark blue arrow (↓) indicates the site of attachment of the spin label. Positive values indicate larger PRE effects in the acid unfolded state. Data were collected at 25ºC, with the pH adjusted to 2.0, by adding HCl, for the DSE in acid and in 8 M urea at pH 2.5 for the DSE in urea.

2.5. References:

[1] Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein, *Proc Natl Acad Sci USA 105*, 13367-13372.

[2] Baldwin, R. L. (2002) A new perspective on unfolded proteins, *Adv Protein Chem 62*, 361-367.

[3] Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability, *FASEB J 10*, 27-34.

[4] Choy, W. Y., Mulder, F. A. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and Kay, L. E. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques, *J Mol Biol 316*, 101-112.

[5] Bowler, B. E. (2007) Thermodynamics of protein denatured states, *Mol BioSyst 3*, 88-99.

[6] 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.

[7] Anil, B., Song, B. B., Tang, Y. F., and Raleigh, D. P. (2004) Exploiting the right side of the ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability, *J Am Chem Soc 126*, 13194-13195.

[8] Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways, *Arch Biochem Biophys 469*, 100-117.

[9] Polverino de Laureto, 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.

[10] Pavitt, G. D., and Ron, D. (2012) New insights into translational regulation in the endoplasmic reticulum unfolded protein response, *Cold Spring Harbor Perspect Biol 4*, DOI: 10.1101/cshperspect.a012278.

[11] Ma, Y., and Hendershot, L. M. (2004) The role of the unfolded protein response in tumour development: Friend or foe?, *Nat Rev Cancer 4*, 966-977.

[12] 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.

[13] Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat Rev Mol Cell Bio 6*, 197-208.

[14] Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics, *Protein Sci 11*, 739-756.

[15] Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays*, Wiley, New York.

[16] 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., Hasan, 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.

[17] Millett, 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 62*, 241-262.

[18] Candotti, M., Esteban-Martin, S., Salvatella, X., and Orozco, M. (2013) Toward an atomistic description of the urea-denatured state of proteins, *Proc Natl Acad Sci USA 110*, 5933-5938.

[19] Huang, J. R., Gabel, F., Jensen, M. R., Grzesiek, S., and Blackledge, M. (2012) Sequence-specific mapping of the interaction between urea and unfolded ubiquitin from ensemble analysis of NMR and small angle scattering data, *J Am Chem Soc 134*, 4429-4436.

[20] Voets, I. K., Cruz, W. A., Moitzi, C., Lindner, P., Areas, E. P. G., and Schurtenberger, P. (2010) DMSO-induced denaturation of hen egg white lysozyme, *J Phys Chem B 114*, 11875-11883.

[21] Tanford, C., Kawahara, K., and Lapanje, S. (1966) Proteins in 6 M guanidine hydrochloride. Demonstration of random coil behavior, *J Biol Chem 241*, 1921-1923.

[22] Chen, L. L., Wildegger, G., Kiefhaber, T., Hodgson, K. O., and Doniach, S. (1998) Kinetics of lysozyme refolding: Structural characterization of a non-specifically collapsed state using time-resolved X-ray scattering, *J Mol Biol 276*, 225-237.

[23] Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell Universtiy Press, Ithaca, N. Y.

[24] Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways, *Arch Biochem Biophys 469*, 100-117.

[25] Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, *J Mol Biol 224*, 771-782.

[26] 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.

[27] Tollinger M, Crowhurst, K. A., Kay, L. E., and Forman-Kay, J. D. (2003) Site-specific contributions to the pH dependence of protein stability, *Proc Natl Acad Sci USA 100*, 4545-4550.

[28] Kuhlman, B., Luisi, D. L., Young, P., and Raleigh, D. P. (1999) pKa values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions, *Biochemistry 38*, 4896-4903.

[29] Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *J Mol Graphics Modell 19*, 26-59.

[30] 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.

[31] 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.

[32] Swint-Kruse, L., and Robertson, A. D. (1995) Hydrogen bonds and the pH dependence of ovomucoid third domain stability, *Biochemistry 34*, 4724-4732.

[33] 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.

[34] 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.

[35] 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.

[36] Sung, Y. H., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family, *J Biol Chem 372*, 689-707.

[37] 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.

[38] Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318*, 571-582.

[39] Horng, J. C., Cho, J. H., and Raleigh, D. P. (2005) Analysis of the pH-dependent folding and stability of histidine point mutants allows characterization of the denatured state and transition state for protein folding, *J Mol Biol 345*, 163-173.

[40] Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2007) 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.

[41] 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.

[42] Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9, *J Mol Biol 349*, 839-846.

[43] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36*, 1277-1282.

[44] Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering, *J Am Chem Soc 129*, 5656-5664.

[45] Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol 278*, 313-352.

[46] Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processingsystem based on Unix Pipes, *J Biomol NMR 6*, 277-293.

[47] Vitalis, A., and Pappu, R. V. (2009) ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions, *J Comput Chem 30*, 673-699.

[48] Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *J Phys Chem B 105*, 6474-6487.

[49] Meng, W., Lyle, N., Luan, B., Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure, *Proc Natl Acad Sci USA 110*, 2123-2128.

[50] 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.

[51] 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.

[52] Luan, B., Shan, B., Baiz, C., Tokmakoff, A., and Raleigh, D. P. (2013) Cooperative cold denaturation: The case of the C-terminal domain of ribosomal protein L9, *Biochemistry 52*, 2402-2409.

[53] Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure, *J Am Chem Soc 132*, 4669-4677.

[54] Manning, M. C., and Woody, R. W. (1991) Theoretical Cd studies of polypeptide helices: Examination of important electronic and geometric factors, *Biopolymers 31*, 569-586.

[55] Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins, *J Am Chem Soc 132*, 8407-8418.

[56] Xue, Y., Podkorytov, I. S., Rao, D. K., Benjamin, N., Sun, H. L., and Skrynnikov, N. R. (2009) Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain, *Protein Sci 18*, 1401-1424.

[57] Perez, Y., Gairi, M., Pons, M., and Bernado, P. (2009) Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: Insights into the role of phosphorylation of the unique domain, *J Mol Biol 391*, 136-148.

[58] Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009) Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP, *Biochemistry 48*, 2115-2124.

[59] Franzmann, M., Otzen, D., and Wimmer, R. (2009) Quantitative use of parramagnetic relaxation enhancements for determining orientations and insertion depths of peptides in micelles, *Chem Bio Chem 10*, 2339-2347.

[60] Allison, J. R., Varnai, P., Dobson, C. M., and Vendruscolo, M. (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements, *J Am Chem Soc 131*, 18314-18326.

[61] 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.

[62] Iwahara, J., and Clore, G. M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR, *Nature 440*, 1227-1230.

[63] 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.

[64] Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations, *J Am Chem Soc 127*, 476-477.

[65] Lindorff-Larsen, 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.

[66] Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, *J Mol Biol 324*, 349-357.

[67] 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.

[68] Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease .2. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures, *J Mol Biol 268*, 170-184.

[69] Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease .1. Paramagnetic relaxation enhancement by nitroxide spin labels, *J Mol Biol 268*, 158-169.

[70] Kosen, P. A., Scheek, R. M., Naderi, H., Basus, V. J., Manogaran, S., Schmidt, P. G., Oppenheimer, N. J., and Kuntz, I. D. (1986) Two-dimensional ^{1}H NMR of three spin-labeled derivatives of bovine pancreatic trypsin inhibitor, *Biochemistry 25*, 2356-2364.

[71] Zhou, H. X. (2004) Polymer models of protein stability, folding, and interactions, *Biochemistry 43*, 2141-2154.

[72] Bernado, P., and Blackledge, M. (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering, *Biophys J 97*, 2839-2845.

[73] Tran, H. T., and Pappu, R. V. (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions, *Biophys J 91*, 1868-1886.

[74] Tran, H. T., Wang, X. L., and Pappu, R. V. (2005) Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins, *Biochemistry 44*, 11369-11380.

[75] 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.

[76] Ding, F., Jha, R. K., and Dokholyan, N. V. (2005) Scaling behavior and structure of denatured proteins, *Structure 13*, 1047-1054.

[77] Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc Natl Acad Sci USA 101*, 12497-12502.

[78] 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.

[79] Yoo, T. Y., Meisburger, S. P., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., and Plaxco, K. (2012) Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state, *J Mol Biol 418*, 226-236.

[80] 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.

[81] Wirmer, J., Peti, W., and Schwalbe, H. (2006) Motional properties of unfolded ubiquitin: A model for a random coil protein, *J Biomol NMR 35*, 175-186.

[82] Das, R. K., and Pappu, R. V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues, *Proc Natl Acad Sci USA 110*, 13392-13397.

Chapter 3

Analysis of Denatured State Ensembles under Conditions Where the Folded Ensemble is Also Populated: Equilibrium SAXS Studies of the CTL9 Denatured State Ensembles

Abstract

Studies of the structural properties of the DSE under conditions where it co-exists with the native state are of great interest, because they provide clues about the features of the DSE under near native conditions. In this chapter, I describe small angle X-ray scattering (SAXS) studies of CTL9 DSSE conducted as a function of pH and urea. The folded and unfolded populations of CTL9 are dependent on pH (or urea concentration) in acid (or urea) induced denaturation, and the folding transition is a two-state process for both cases. This feature allows us to study the CTL9 DSE that is co-populated with the folded state. Between pH 2.0 and 2.5, the protein is essentially 100% in the DSE. The population of the folded CTL9 begins to increase starting at pH 2.5. At pH 3.8, ~ 50% of the population is folded and ~ 50% is in the DSE. Above pH 5.5, CTL9 becomes essentially 100% folded. Here, the acid-induced CTL9 DSE was studied by SAXS in buffer over the pH range from 2.0 to 3.8. The acid-induced CTL9 DSE at pH 2.0 is expanded: it has a similar value of R_g as that predicted by random coil scaling behavior. By de-convolving the contribution from folded CTL9, the CTL9 DSE scattering pattern can be obtained between pH 2.5 and 3.8. The SAXS data demonstrated that the R_g of the acid-induced CTL9 DSE undergoes significant compaction as the pH increases. The CTL9 DSE populated by moderate amounts of urea was also studied. The population of the urea-induced CTL9 DSE is ~ 50% in 4 M urea at pH 5.5. SAXS analysis showed that the urea-

induced DSE populated under these conditions is as expanded as the fully denatured urea-induced DSE.

3.1. Introduction

Studies of the structural properties of the protein denatured state ensemble (DSE) under conditions where it co-exists with the native state are of great interest, because they provide clues about the features of the DSE under near native conditions. Protein structure is highly dependent on the solvent or the environment. Ionic strength, pH, and the concentration of chemical denaturant are examples of parameters that can affect protein structure in aqueous solution.

The protonation and de-protonation of biomolecules in aqueous solution are fundamental mechanisms to modulate ionization protein states and their physicochemical properties, thus they are of great relevance for biological systems. It has been reported that pH values between 4.0 and 5.5 in the endosomes and lysosomes in eukaryotic cells, while pH ranging from 7.2 to 7.4 in the cytosol.[1] Given the small differences in the stability between protein folded state and the DSE, small changes in the structure disrupting factors, such as pH, may lead to substantial and even cooperative changes in the population of the protein folded and unfolded state, respectively.

Urea is one of the most frequently used chemicals to populate protein denatured ensembles (DSE) in water.[2] Proteins that are dissolved in high concentration of urea have flexible conformations that have been found to resemble to the concept of the random coil model. Despite the fact that long range interactions and residual structures have been found for proteins in high concentrations of urea, R_g values often agree well with the

scaling law for an extended polymer in a "good solvent".[3-6] The mechanism of protein denaturation induced by urea have been leaned towards the direct interaction theory, such as protein backbone, side-chain, or a mixture of polar and hydrophobic residues.[7-14] Urea forms hydrogen bonds with the protein backbone, based on the detailed analysis of the population of urea, it has been revealed that apolar residues are the major targets for this reaction.[15] In the urea-induced protein unfolding, it is the dispersion caused by the urea-specific solvation that contributes to the stabilization of the protein denatured ensembles and the irreversible unfolding process in the presence of urea.[15]

The dimension and long range interactions within the low-pH and high urea concentration induced DSE of C-terminal domain of the ribosomal protein L9 (CTL9) have been studied and described previously (Chapter 2). In order to obtain information on the dimensional properties or the compactness of CTL9 DSE in the milder conditions, such as intermediate pH values and lower concentrations of urea, equilibrium small angle X-ray scattering (SAXS) were used and the results are summarized in this chapter. Together with time-resolved SAXS experiments in the future studies, the protein DSE populated with folded ensembles at mild denaturing conditions can be examined in a more detailed manner.

3.2. Materials and methods

3.2.1. Protein expression and purification

Wild type CTL9 was over expressed in the *Escherichia coli* strain BL21 cells in LB media and purified as described previously.[16] The cells were grown at 37 °C until the optical density at 600 nm reached 0.8-0.9, followed by 1 mM IPTG (isopropylthio-β-D-galactoside) induction for 4 hours at 37 °C. Cells were harvest and lysed by sonication, and the lysate was centrifuged for 1 hr at 10,000 g to remove the cell debris. The supernatant was loaded onto a SP-Sepharose fast flow ion-exchange column (GE Healthcare). 20 mM Tris buffer at pH 7.4 was used as the equilibration buffer, and the protein was eluted with a 0 to 2 M NaCl gradient. Proteins were further purified with the reverse phase HPLC using a C8 column (Vydac). An A-B gradient was created using buffer A containing of 99.9% H_2O and 0.1% TFA (trifluoroacetic acid) and buffer B containing of 90% acetonitrile, 9.9% H_2O, and 0.1% TFA. The protocol used is 100% A for 10 mins after injecting the sample followed by 25-55% B in 60 mins. CTL9 eluted around 40-42% B. The yield of wild type CTL9 was ~ 70 mg/L. The identity of the protein was confirmed by MALDI-TOF mass spectroscopy.

3.2.2. Circular dichroism (CD) spectroscopy

A pH titration curve was measured for wild type CTL9 at 25°C, in 10 mM sodium acetate and 150 mM NaCl, over the range of pH 1.8 to 12.0. Urea unfolding data were collected at 25°C on wild type CTL9, using an AVIV Instruments model 202SF CD instrument. The native protein sample was titrated with urea in 0.25 M increments until the final urea concentration reached 10 M. The buffer was the same as used for the pH titration CD

experiments. The urea concentration was determined by measuring the refractive index. The concentration of the protein is ~ 19 μM for both cases.

3.2.3. Small angle X-ray scattering measurements

Samples of wild type CTL9 were prepared in buffer consisting of 10 mM sodium acetate and 150 mM NaCl, with pH adjusted to the range from 2.0 to 3.8 for the acid-induced DSE experiments. Protein samples containing urea also used this buffer with 10 mM sodium acetate and 150 mM NaCl. The urea concentration was determined by measuring the refractive index.

Scattering experiments were performed at beamline X9 at Brookhaven National Laboratory, National Synchrotron Light Source I (Upton, New York, USA). Protein samples were injected into a 1 mm capillary continuously during the measurement at a rate of 0.67 μL/s in order to avoid radiation damage. The exposure time for each measurement was 30 s. Scattering data was collected for wild type CTL9 at a protein concentration of 3.75 mg/mL, at 25 °C. Each sample was measured three times and then averaged before data analysis.

Guinier analysis was directly performed on the protein samples that were fully unfolded. For samples that populated both the folded state and DSE, the scattering profiles of the CTL9 DSE were obtained after subtracting the population-weighted folded signal, using:

$$I_{\mathrm{CTL9,\ DSE}} = (I_{\mathrm{CTL9,\ observed}} - p_{\mathrm{folded}} \cdot I_{\mathrm{CTL9,\ folded}}) / (1 - p_{\mathrm{folded}}) \qquad \text{(eq 3-1)}$$

where $I_{CTL9,\ observed}$ is the SAXS data collected for wild type CTL9 containing both the folded and unfolded ensembles; $I_{CTL9,\ folded}$ is the scattering profile collected for the wild type CTL9 at pH 7.0, only folded state is populated under this condition, providing a folded scattering profile for the subtraction; p_{folded} is the population of folded CTL9 present under the conditions of the SAXS measurements, as determined by the pH or urea titration monitored by CD; $I_{CTL9,\ DSE}$ is the calculated CTL9 DSE scattering profile.

The program pyXS (http://www.bnl.gov/ps/x9/software/pyXS.asp) was used for buffer subtraction. The radius of gyration (R_g) was obtained based on the scattering profiles using the Guinier approximation using the program PRIMUS,[17]

$$I(q) = I(0) \bullet \exp(-R_g^2 q^{2/3}) \qquad \text{(eq 3-2)}$$

where $I(q)$ is the intensity at scattering vector q.[18]

In addition, a Guinier analysis was done directly on CTL9 scattering data in order to obtain the apparent R_g values, $R_{g,\ app}$, at each temperature point. The value of the R_g for the DSE, $R_{g,\ DSE}$ can then be estimated from:

$$R_g^2{}_{app} = p_{folded} \cdot R_g^2{}_{folded} + p_{DSE} \cdot R_g^2{}_{DSE} \qquad \text{(eq 3-3)}$$

where p_{folded} and p_{DSE} are the fractional populations of the folded ensemble and the DSE, respectively.

3.3. Results

3.3.1. CTL9 follows a two state unfolding transition as a function of pH

There are three His in the wild type CTL9 primary sequence, making the protein subject to the pH-dependent folding and unfolding in the absence of chemical denaturant. The apparent two-state acid-induced unfolding process of CTL9 has been characterized by following the ellipticity at 222 nm monitored by CD (**Figure 3-1**). Below pH 3, the DSE is populated almost exclusively, while above pH 5.0, only the folded form is populated as judged by CD. At pH values between 3 and 5.0, both the folded CTL9 and acid-induced DSE are significantly populated at the same time. Based on the folded and unfolded baseline in the high and low pH regime, respectively, the fraction folded can be estimated for pH values between 3 and 5.0.

3.3.2. The acid induced CTL9 is less expanded under more native conditions

SAXS data were collected on wild type CTL9 over the pH range from 2.0 to 3.8, and a control data set at pH 7.0 was also collected to provide a folded state scattering pattern for data analysis (**Figure 3-2A**). The existence of an isosbestic point in the scattering profile is consistent with a two-state process for the acid denaturation of CTL9. Scattering curves at pH 2.0, 2.4, and 2.7 are very similar to each other, and at those pH values, the fraction fold is essentially 0. As the pH increases from 3.0 to 3.8, the scattering curves become more similar to the folded scattering data collected at pH 7.0, indicating that the population of the folded state is increasing.

In order to obtain information on the size of the acid-induced CTL9 DSE under conditions where there is a substantial population of the folded ensemble, the scattering

data at pH values from 3.0 to 3.8 were analyzed by subtracting the population weighted folded signal from the observed scattering data (**Figure 3-2B**), the resulting difference signal represents the scattering profile of the DSE. A Guinier analysis of the scattering profiles of the CTL9 DSE between pH 2.0 and 3.8, together with the folded control at pH 7.0 is shown in **Figure 3-3**. The value of R_g of the acid-induced CTL9 DSE in the pH range from 2.0 to 3.8 is plotted in **Figure 3-4(A)**, and the value of R_g of the acid-induced CTL9 DSE was also plotted as a function of the estimated net charge on the protein in **Figure 3-4(B)**. CTL9 contains 3 His, 3 Arg, 12 Lys, 2 Asp, and 9 Glu. The net charge was calculated using model pKa values of 3.9 for Asp, 4.3 for Glu, 6.8 for His, 10.4 for Lys, 12.3 for Arg, 8.0 for the N-terminus and 3.0 for the C-terminus. The plot exhibits a clear trend of compaction, the R_g values decrease as pH increases, and at pH 3.8, where the fraction folded is 0.5, the R_g of the acid-induced CTL9 DSE shrinks by ~ 35 % to a value of 18 Å, compared with the fully unfolded R_g of 28 Å. We also calculated the R_g values of the DSE based on the equation of the two component mixture (eq 3-3). The fraction of fold at each pH and the calculated R_g of the CTL9 DSE are listed in **Table 3-1**.

3.3.3. The urea induced DSE of CTL9 is expanded over a wide range of pH values

SAXS data was also collected on wild type CTL9 in 10 M urea over the pH range between 2.5 and 5.5 (**Figure 3-5A**). The solution contains the same buffer as used in the acid-induced CTL9 SAXS studies to maintain constant ionic strength. In 10 M urea, CTL9 is fully unfolded over the pH range of this study as judged by CD. The urea

induced DSE is expanded over the pH range studied with an R_g value of 27.8 ± 1.1 Å at pH 2.5 and 24.3 ± 0.6 Å at pH 5.5 (**Figure 3-5B**, **Figure 3-6**). The R_g values were also plotted as a function of net charge per residue in **Figure 3-7**.

3.3.4. The CTL9 DSE populated in low concentration urea is expanded

The dimension of the urea-induced CTL9 DSE as a function of urea concentration was examined from 4 to 10 M urea. As shown by a urea titration on wild type CTL9 monitored by CD, at pH 5.5, a transition from the folded ensemble to the fully unfolded DSE is observed (**Figure 3-8**), with a midpoint of 3.9 M urea. At pH 5.5, for urea concentrations below 2 M, the fraction folded of CTL9 is essentially 1, while at urea concentration higher than 5 M, CTL9 fully populates the DSE. The fraction folded at 4 M urea is calculated to be 0.4. A urea titration conducted at pH 2.5 shows that CTL9 is fully unfolded for all concentrations between 0 and 10 M.

Scattering data were collected on the urea-induced CTL9 DSE at urea concentration from 4 M to 10 M. Guinier analysis of the scattering profile was performed directly on the samples with urea concentration higher than 4 M (**Figure 3-9**), since CTL9 is fully unfolded under those conditions. For the sample containing 4 M urea, the population weighted contribution from the folded state was subtracted from the observed scattering data (**Figure 3-10A**), followed by the Guinier analysis on the urea-induced CTL9 DSE scattering data (**Figure 3-10B**).

The R_g values of the urea-induce CTL9 DSE at pH 5.5 are plotted as a function of urea concentration over the range of 4-10 M urea in **Figure 3-11**. There is no obvious change in R_g as the urea concentration decreases, with values ranging from 26.3 ± 1.4 Å to 30.8 ± 1.6 Å.

3.4. Discussion

The CTL9 DSE populated in the absence of urea shows a clear trend of compaction with decreasing charge on the protein. In contrast, no significant change in R_g was observed as a function of urea, or as a function of pH for samples in 10 M urea.

In the pH-dependent CTL9 DSE experiments, as the pH value increases from 2.5 to 3.8, the population of the fraction folded increases from 0 to ~0.5. The DSE undergoes a compaction towards a more native pH values as the fraction folded increases. Similar pH-induced structural transformations have been reported on intrinsically disordered proteins (IDPs), in which partial folding of extended IDPs has been observed with changes in pH, as judged by the changes in the far-UV CD spectrum.[19-22] This could be rationalized by the decrease of the net charge present at a more neutral pH compared with low pH values.[23] Charge-charge intramolecular repulsion is decreased correspondingly, permitting collapse driven by hydrophobic interactions, thus the DSE populated at lower net charge per residue is expected to be more compact.[24]

Previously, several sets of secondary chemical shift analysis based on the deviation from the random coil model by NMR studies has been performed on wild type CTL9 DSE at different pH values.[25, 26] It has been reported that at pH 3.8, the DSE shows a trend to form α-helices in the two segments corresponding to the helices in the native. Furthermore, compared with the acid-induced CTL9 DSE at pH 2.0, the DSE populated at pH 3.8 has noticeably higher propensity to populate secondary structure.[25] Hydrodynamic measurements demonstrated that the pH 2.0 unfolded state is more expanded than the pH 3.8 unfolded state.[27] Both of these two findings agree well with the results obtained by equilibrium SAXS here. In contrast to the studies in the absence of urea, there is a much weaker change in R_g versus net charge in 10 M urea.

There is no detectable change in the R_g of the CTL9 DSE as a function of urea at pH 5.5, which in good agreement with results on protein L.[28] Similar results have also been reported on other proteins studied by equilibrium SAXS measurements, and have demonstrated that the dimensions of the unfolded ensembles are independent of denaturant concentration.[29-33] This may be explained by urea being a good solvent for proteins, even at moderate concentrations. Urea forms hydrogen bond with the protein backbone and side-chain, further extend the protein into "random coil" like states. Time resolved stopped flow SAXS studies agree that for same single domain proteins, the dimensions of the transient state, which is formed at low concentration of denaturant after rapid dilution, cannot be distinguished from those observed at higher denaturant concentration.[34-37] In contrast, studies using time-resolved smFRET argue that compaction happens as it is diluted from high concentration of urea. The disagreement

between SAXS and smFRET in studies of the protein low-denaturant unfolded state is far from reconciliation.

Table 3-1. R_g values, net charge, and fraction folded of wild type CTL9 at pH range from 3 to 4.

pH	Estimated net charge of the DSE[a]	Fraction folded[b]	$R_{g,\ DSE}$ (Å)[c]	$R_{g,\ DSE}$ (Å)[d]
3.06	16.7	0.1	24.3 ± 1.2	22.9 ± 1.8
3.2	16.4	0.13	22.4 ± 1.2	21.2 ± 1.7
3.35	16	0.2	23.0 ± 1.3	21.1 ± 1.5
3.62	15	0.44	20.5 ± 1.2	20.3 ± 1.4
3.8	14.1	0.63	18.4 ± 1.7	19.6 ± 1.5

a) Estimated net charge at given pH, using the model pKa's listed in the text; *b*) Fraction of folded CTL9 estimated from the pH titration monitored by CD; *c*) R_g values of the CTL9 acid unfolded DSE obtained by Guinier analysis on the DSE scattering profile after subtracting the signal from the weighted folded state contribution; *d*) R_g values of the CTL9 acid unfolded DSE calculated using the equation for a two-component mixture.

Table 3-2. The R_g value, net charge, and fraction folded of wild type CTL9 at pH 5.5 in urea.

[urea] (M)	Estimated net charge of the DSE[a]	Fraction folded[b]	$R_{g,\ DSE}$ (Å)[c]	$R_{g,\ DSE}$ (Å)[d]
4	6.4	0.4	29.2 ± 1.6	33.0 ± 1.2
5	6.4	0	27.2 ± 1.2	N/A
6	6.4	0	29.4 ± 1.3	N/A
7	6.4	0	26.3 ± 1.4	N/A
8	6.4	0	30.8 ± 1.6	N/A
9	6.4	0	28.7 ± 1.5	N/A
10	6.4	0	27.7 ± 1.6	N/A

a) Estimated net charge at pH 5.5, using the model pKa's listed in the text; *b*) Fraction of folded CTL9 estimated from the urea titration monitored by CD, the protein appears fully unfolded at 5 M urea and above; *c*) R_g values of the CTL9 urea unfolded DSE obtained by Guinier analysis on the DSE scattering profile after subtracting the signal from the weighted folded state contribution; *d*) R_g values of the CTL9 urea unfolded DSE calculated using the equation for a two-component mixture.

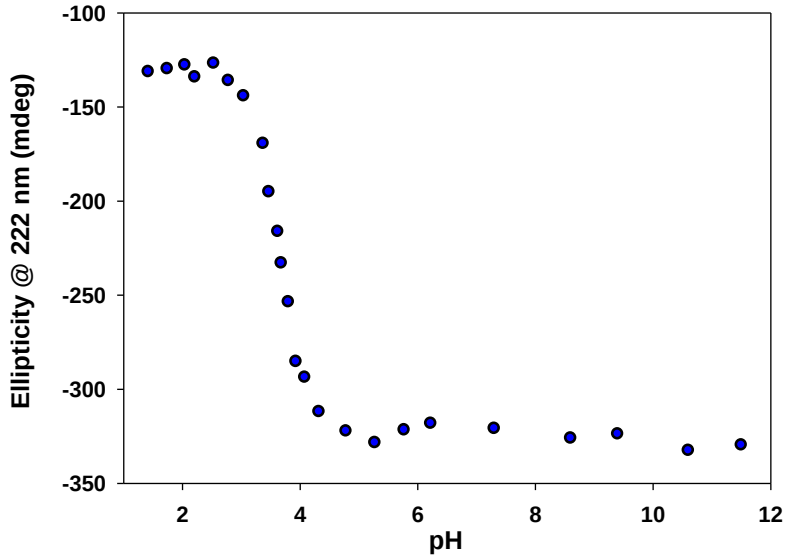
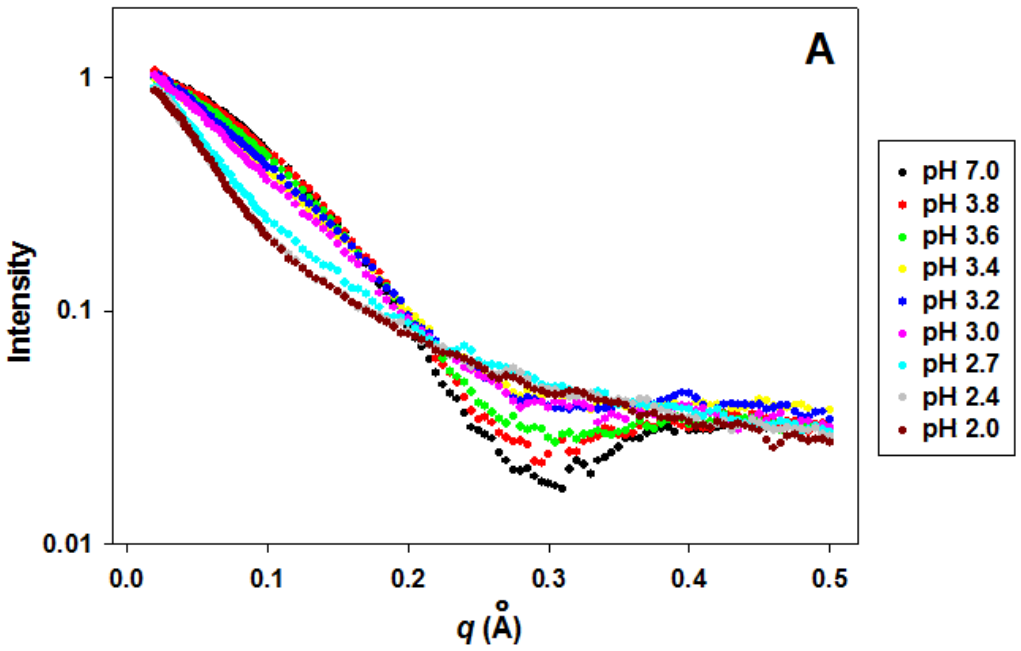


Figure 3-1. A pH titration of wild type CTL9 monitored by CD at 222 nm. The buffer contains 10 mM sodium acetate and 150 mM NaCl. The protein concentration is 19 μM. The transition is two state and the fraction folded can be determined at a given pH.


A
Intensity
1
0.1
0.01
0.0
0.1
0.2
0.3
0.4
0.5
q (Å)
pH 7.0
pH 3.8
pH 3.6
pH 3.4
pH 3.2
pH 3.0
pH 2.7
pH 2.4
pH 2.0

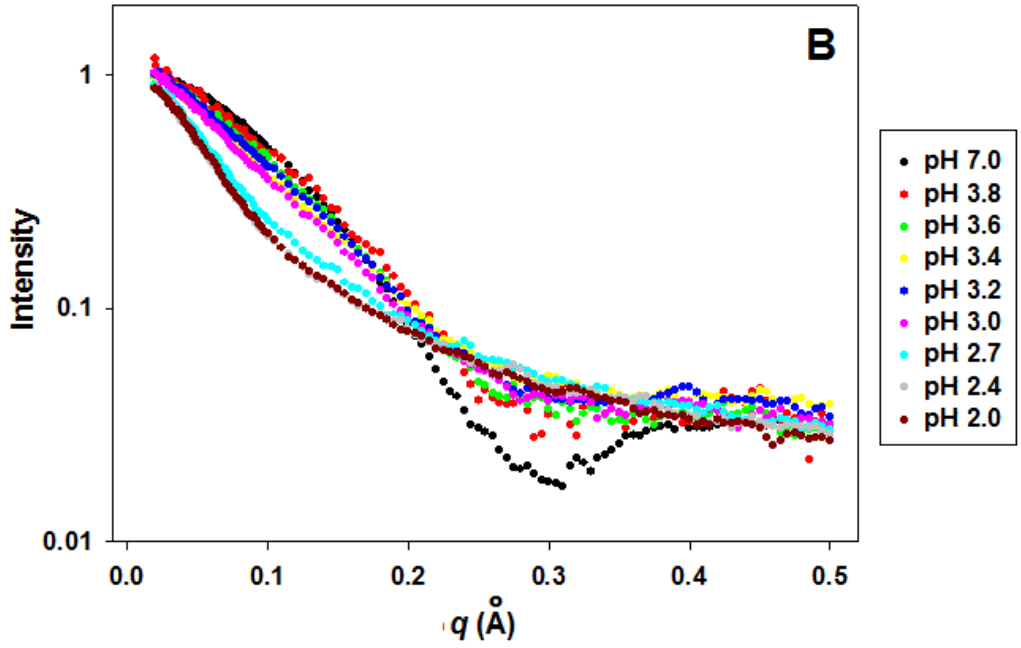

B
Intensity
1
0.1
0.01
0.0
0.1
0.2
0.3
0.4
0.5
q (Å)
pH 7.0
pH 3.8
pH 3.6
pH 3.4
pH 3.2
pH 3.0
pH 2.7
pH 2.4
pH 2.0

Figure 3-2. Scattering profiles for wild type CTL9 at different pH values. (A) Observed scattering profile; (B) Scattering profile for the acid induced DSE after subtracting the population weighted contribution from the folded ensemble (pH values between 3.0 and 3.8).

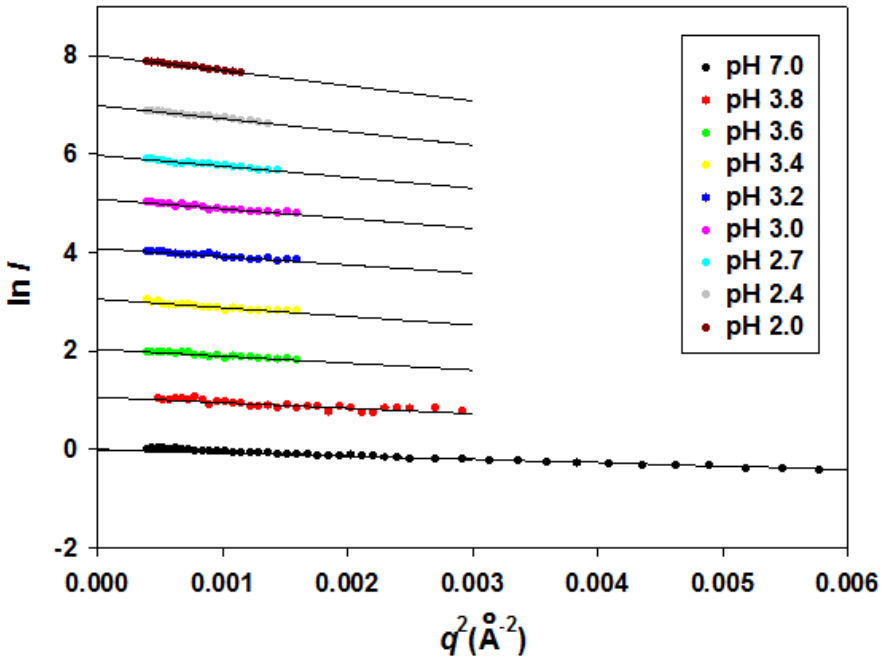


Figure 3-3. Guinier approximation analysis based on the CTL9 DSE scattering profiles in Figure 1(B). The curves are shifted along the y-axis for clarity.

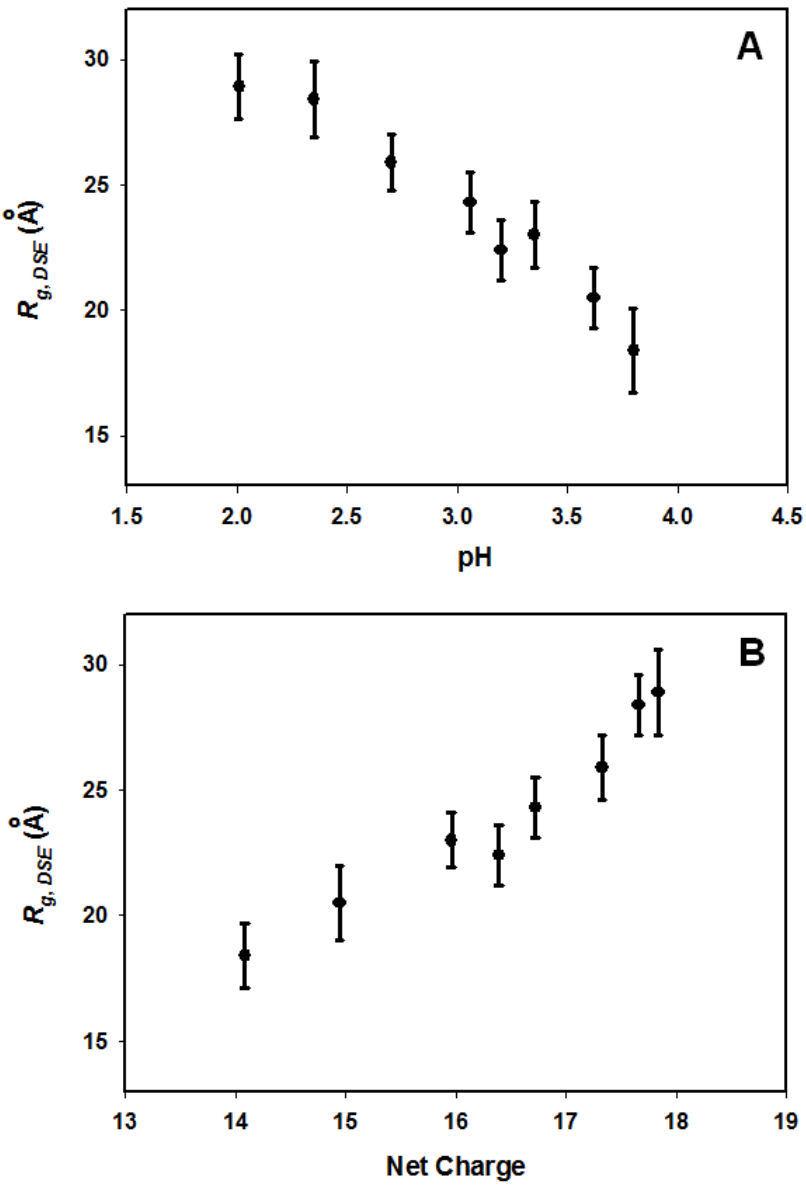


Figure 3-4. (A) Plot of the R_g of the acid-induced CTL9 DSE as a function of pH. The buffer is 10 mM sodium acetate and 150 mM NaCl. R_g was calculated from the difference scattering profile. For reference, the R_g of the folded state is 15 Å, and the value in 10 M urea at pH 2.5 is 28 Å. Error bars were obtained from the Guinier fitting. (B) Plot of R_g versus the estimated net charge of the DSE. The net charge was calculated over the pH range of 2.0 to 3.8 using the model pKa's listed in the text.

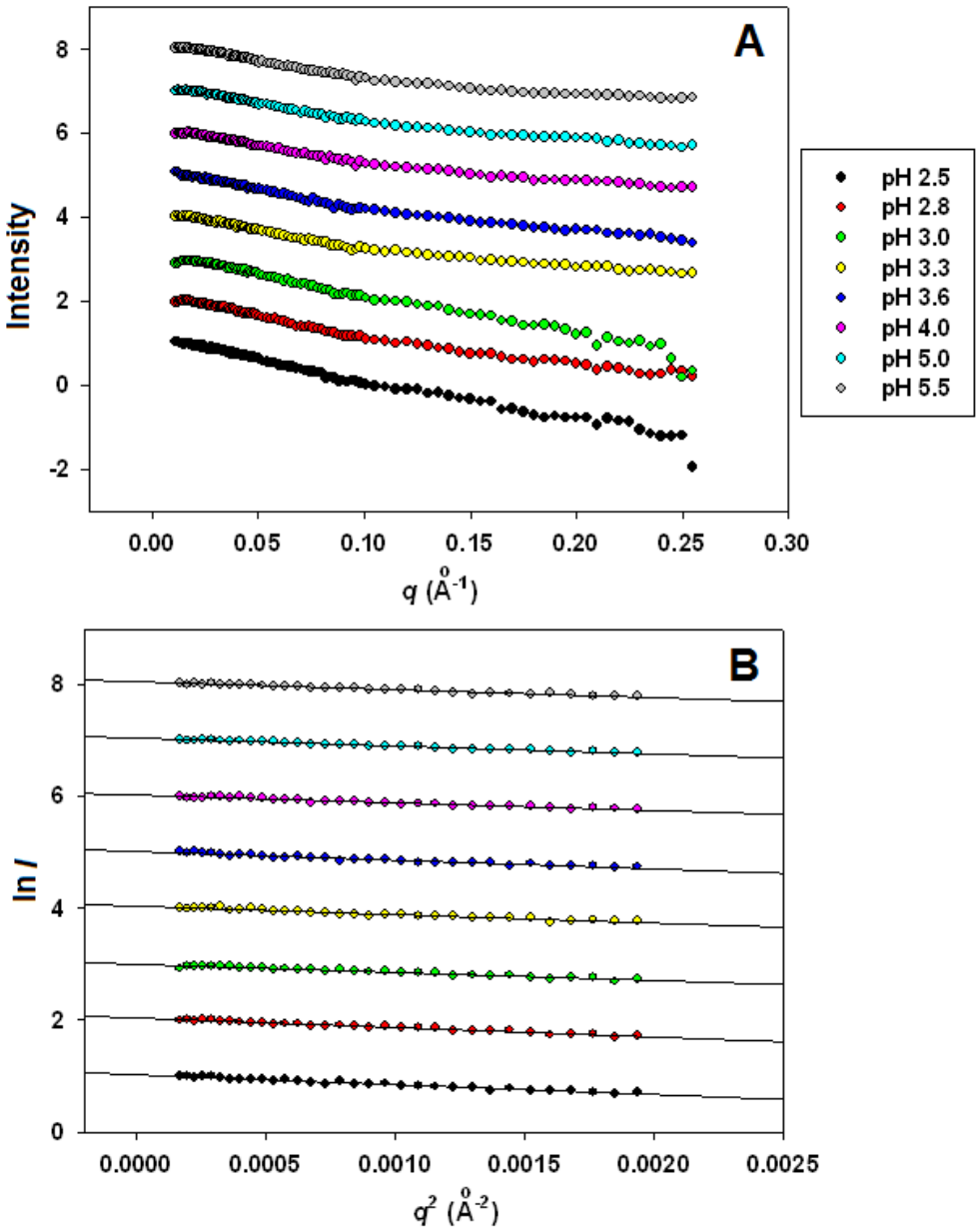


Figure 3-5. (A) Scattering profiles of the urea-induced CTL9 DSE in 10 M urea over the pH range of 2.5 to 5.5. (B) Guinier analysis of the scattering profiles shown in panel (A). The curves are offset for clarity. The buffer contains 10 mM sodium acetate and 150 mM NaCl.

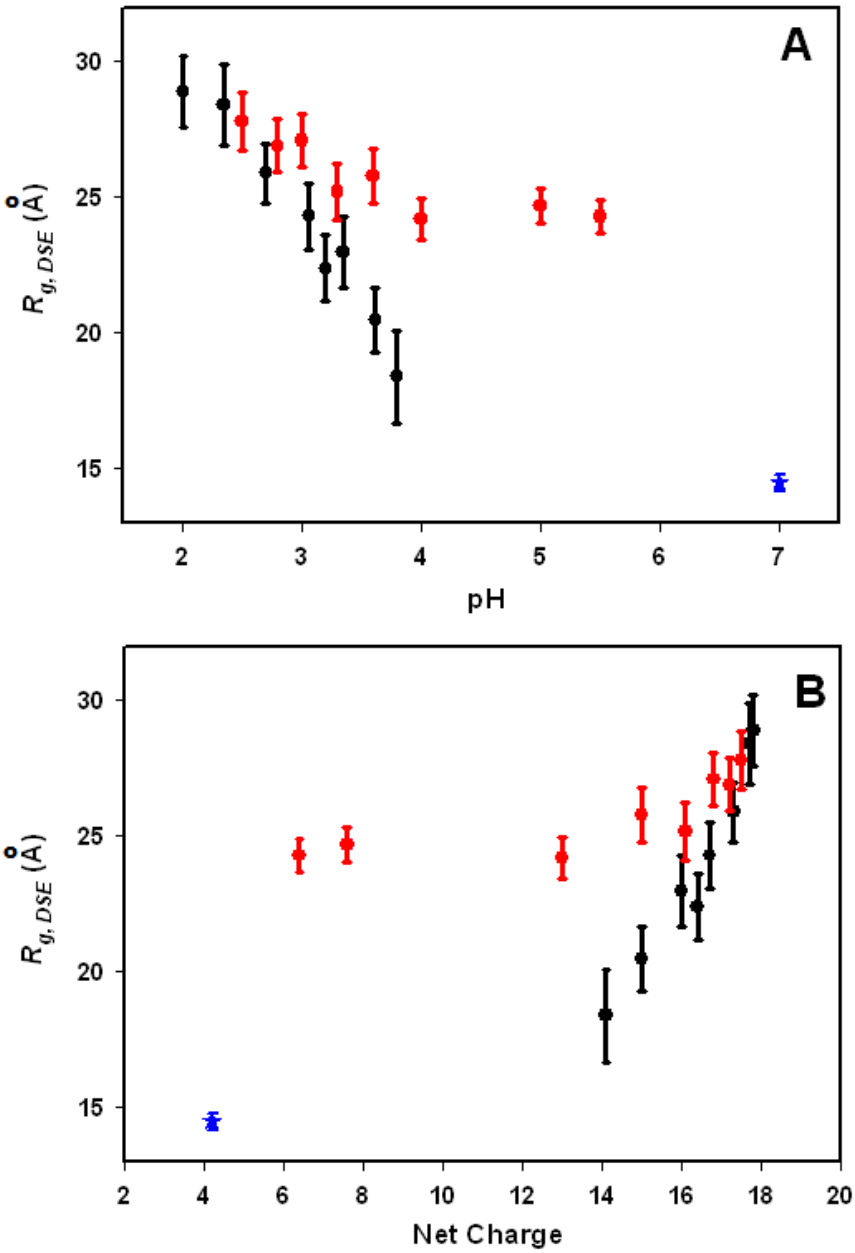


Figure 3-6. (A) A plot comparing the R_g of the 10 M urea-induced CTL9 DSE as a function of pH (red) to the R_g of the CTL9 DSE in buffer (black). R_g of the folded state at pH 7 is shown in blue star as a control. The buffer contains 10 mM sodium acetate and 150 mM NaCl. (B) Comparison of the R_g of the urea and acid unfolded DSEs versus the estimated net charge on the protein.

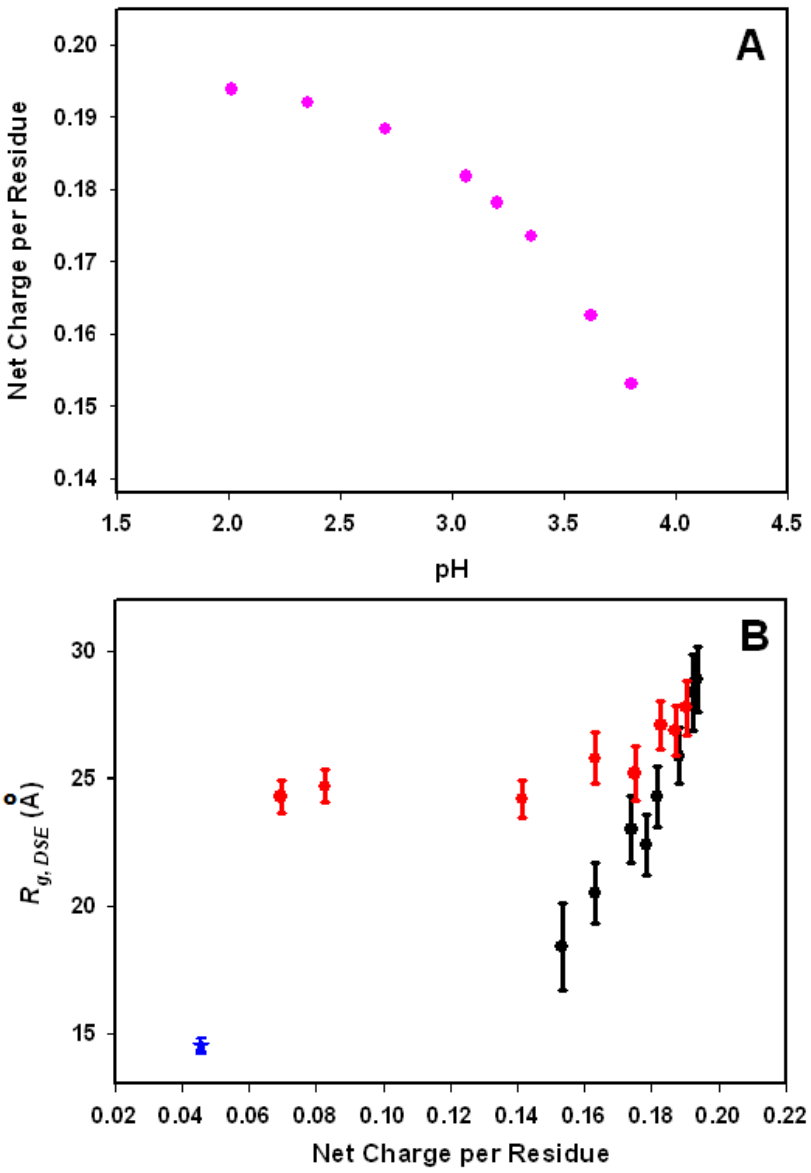


Figure 3-7. (A) Net charge per residue as estimated for different pH values for CTL9. (B) Plot of R_g of the DSE in buffer (black) and the urea unfolded DSE (red) versus net charge per residue. The R_g value of the folded state CTL9 at pH 7 is shown as a control (blue symbol).

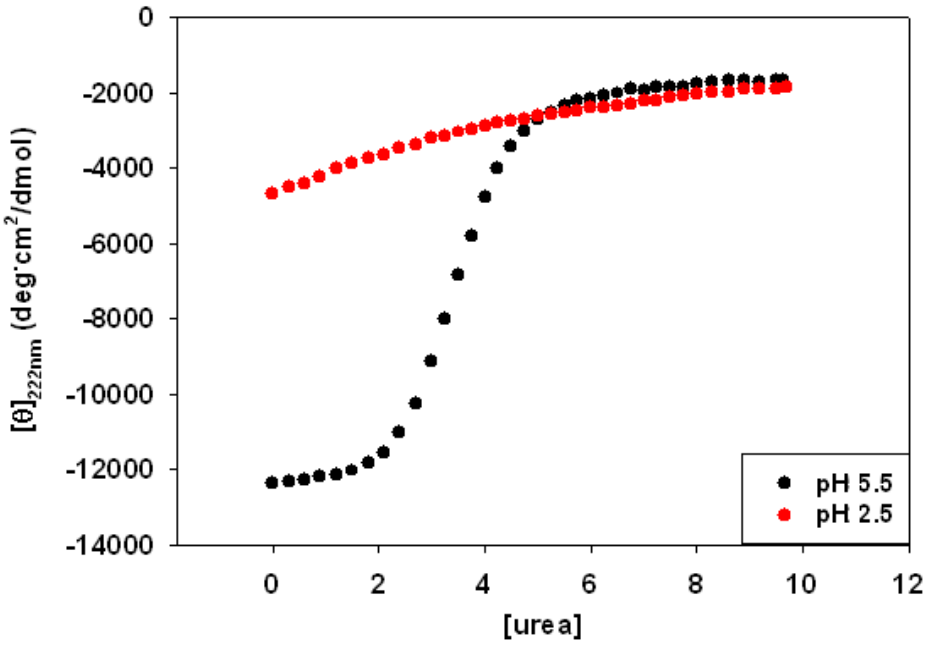


Figure 3-8. The mean residue ellipticity at 222 nm of wild type CTL9 monitored by CD as a function of urea concentration, the buffer consists of 10 mM sodium acetate and 150 mM NaCl. The red curve presents the urea titration experiment carried out at pH 2.5, while the black curve presents the urea titration experiment conducted at pH 5.5. Experiments were conducted at 25 °C.

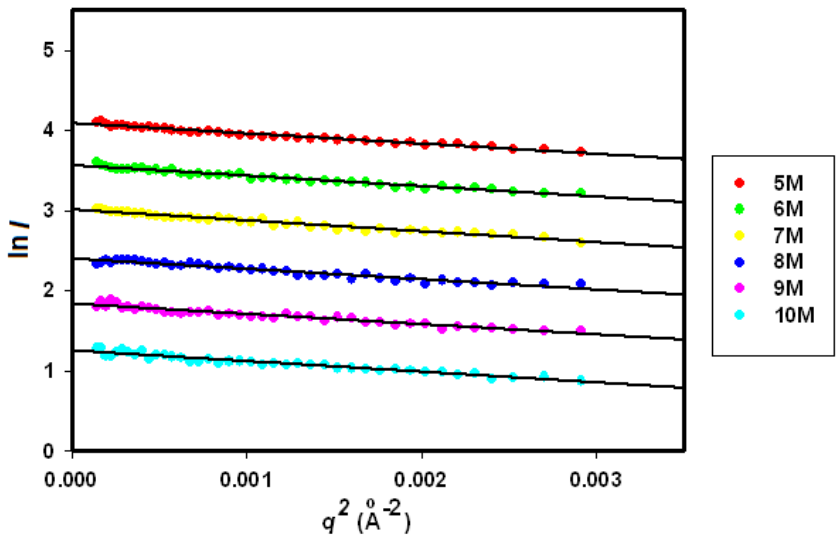


Figure 3-9. Guinier analysis of the scattering profile of wild type CTL9 at different urea concentrations from 5 to 10 M (in which the fraction fold is 0). The buffer contains 10 mM sodium acetate and 150 mM NaCl.

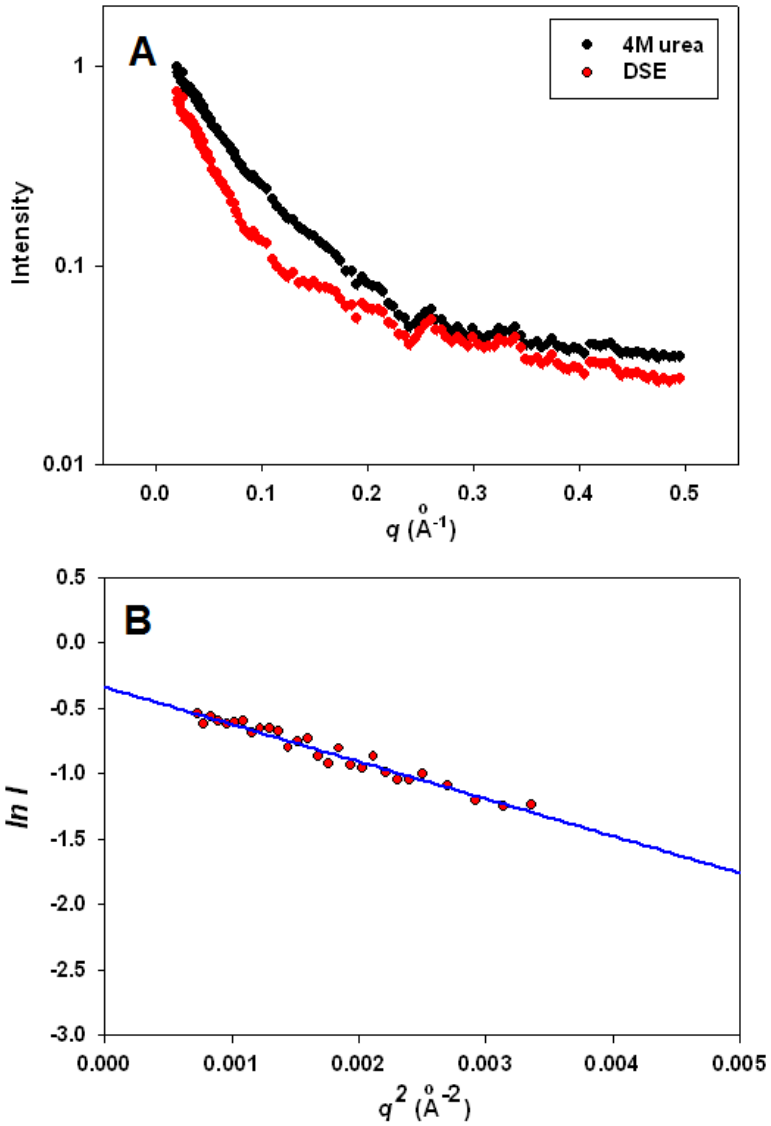


Figure 3-10. (A) Scattering profile for wild type CTL9 in 4 M urea, where the fraction fold is 0.4. The observed scattering curve is shown in black and the difference curve generated by subtracting the population weighted folded signal is shown in red. (B) Guinier analysis on the DSE scattering profile. The buffer contains 10 mM sodium acetate and 150 mM NaCl.

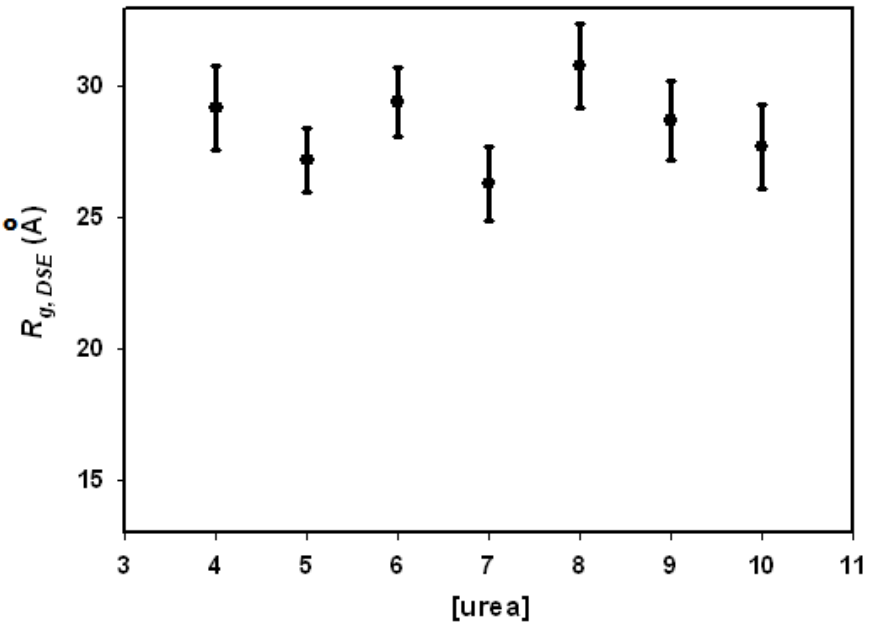


Figure 3-11. Plot of R_g for the urea induced CTL9 DSE at urea concentrations ranging from 4 to 10 M. The buffer contains 10 mM sodium acetate and 150 mM NaCl, at pH 5.5.

3.5. References:

[1] Paroutis, P., Touret, N., and Grinstein, S. (2004) The pH of the secretory pathway: Measurement, determinants, and regulation, *Physiology 19*, 207-215.

[2] England, J. L., and Haran, G. (2011) Role of solvation effects in protein denaturation: From thermodynamics to single molecules and back, *Annu Rev Phys Chem 62*, 257-277.

[3] Luan, B. W., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2014) Denatured state ensembles with the same radii of gyration can form significantly different long-range contacts, *Biochemistry 53*, 39-47.

[4] Meng, W., Lyle, N., Luan, B., Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure, *Proc Natl Acad Sci USA 110*, 2123-2128.

[5] Meng, W., Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2013) The denatured state ensemble contains significant local and long-range structure under native conditions: Analysis of the N-terminal domain of ribosomal protein L9, *Biochemistry 52*, 2662-2671.

[6] 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., Hasan, 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.

[7] Canchi, D. R., and Garcia, A. E. (2011) Backbone and side-chain contributions in protein denaturation by urea, *Biophys J 100*, 1526-1533.

[8] Lim, W. K., Rosgen, J., and Englander, S. W. (2009) Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, *Proc Natl Acad Sci USA 106*, 2595-2600.

[9] Stumpe, M. C., and Grubmuller, H. (2008) Polar or apolar-The role of polarity for urea-induced protein denaturation, *Plos Comput Biol 4*.

[10] Stumpe, M. C., and Grubmuller, H. (2007) Interaction of urea with amino acids: Implications for urea-induced protein denaturation, *J Am Chem Soc 129*, 16126-16131.

[11] O'Brien, E. P., Dima, R. I., Brooks, B., and Thirumalai, D. (2007) Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions: Lessons for protein denaturation mechanism, *J Am Chem Soc 129*, 7346-7353.

[12] Auton, M., Holthauzen, L. M. F., and Bolen, D. W. (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation, *Proc Natl Acad Sci USA 104*, 15317-15322.

[13] Klimov, D. K., Straub, J. E., and Thirumalai, D. (2004) Aqueous urea solution destabilizes A β(16-22) oligomers, *Proc Natl Acad Sci USA 101*, 14760-14765.

[14] Mountain, R. D., and Thirumalai, D. (2003) Molecular dynamics simulations of end-to-end contact formation in hydrocarbon chains in water and aqueous urea solution, *J Am Chem Soc 125*, 1950-1957.

[15] Candotti, M., Esteban-Martin, S., Salvatella, X., and Orozco, M. (2013) Toward an atomistic description of the urea-denatured state of proteins, *Proc Natl Acad Sci USA 110*, 5933-5938.

[16] Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318*, 571-582.

[17] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36*, 1277-1282.

[18] Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays*, Wiley, New York.

[19] Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation, *J Biol Chem 276*, 10737-10744.

[20] Uversky, V. N., Gillespie, J. R., Millett, I. S., Khodyakova, A. V., Vasiliev, A. M., Chernovskaya, T. V., Vasilenko, R. N., Kozlovskaya, G. D., Dolgikh, D. A., Fink, A. L., Doniach, S., and Abramov, V. M. (1999) Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic pH, *Biochemistry 38*, 15009-15016.

[21] Lynn, A., Chandra, S., Malhotra, P., and Chauhan, V. S. (1999) Heme binding and polymerization by Plasmodium falciparum histidine rich protein II: Influence of pH on activity and conformation, *FEBS Lett 459*, 267-271.

[22] Johansson, J., Gudmundsson, G. H., Rottenberg, M. E., Berndt, K. D., and Agerberth, B. (1998) Conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37, *J Biol Chem 273*, 3718-3724.

[23] Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L., and Pappu, R. V. (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins, *Proc Natl Acad Sci USA 107*, 8183-8188.

[24] Uversky, V. N. (2009) Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding, *Protein J 28*, 305-325.

[25] 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.

[26] 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.

[27] Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9, *J Mol Biol 349*, 839-846.

[28] Yoo, T. Y., Meisburger, S. P., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., and Plaxco, K. (2012) Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state, *J Mol Biol 418*, 226-236.

[29] Jacob, J., Dothager, R. S., Thiyagarajan, P., and Sosnick, T. R. (2007) Fully reduced ribonuclease a does not expand at high denaturant concentration or temperature, *J Mol Biol 367*, 609-615.

[30] Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T., Akiyama, S., and Fujisawa, T. (2005) Specific collapse followed by slow hydrogen-bond

formation of β-sheet in the folding of single-chain monellin, *Proc Natl Acad Sci USA 102*, 2748-2753.

[31] Kimura, T., Akiyama, S., Uzawa, T., Ishimori, K., Morishima, I., Fujisawa, T., and Takahashi, S. (2005) Specifically collapsed intermediate in the early stage of the folding of ribonuclease A, *J Mol Biol 350*, 349-362.

[32] Millet, I. S., Townsley, L. E., Chiti, F., Doniach, S., and Plaxco, K. W. (2002) Equilibrium collapse and the kinetic 'foldability' of proteins, *Biochemistry 41*, 321-325.

[33] Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M., and Goto, Y. (1997) Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering, *FEBS Lett 416*, 72-76.

[34] Jacob, J., Krantz, B., Dothager, R. S., Thiyagarajan, P., and Sosnick, T. R. (2004) Early collapse is not an obligate step in protein folding, *J Mol Biol 338*, 369-382.

[35] Konuma, T., Kimura, T., Matsumoto, S., Got, Y., Fujisawa, T., Fersht, A. R., and Takahashi, S. (2011) Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase, *J Mol Biol 405*, 1284-1294.

[36] Uzawa, T., Kimura, T., Ishimori, K., Morishima, I., Matsui, T., Ikeda-Saito, M., Takahashi, S., Akiyama, S., and Fujisawa, T. (2006) Time-resolved small-angle X-ray scattering investigation of the folding dynamics of heme oxygenase: Implication of the scaling relationship for the submillisecond intermediates of protein folding, *J Mol Biol 357*, 997-1008.

[37] Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S., and Baker, D. (1999) Chain collapse can occur concomitantly with the rate-limiting step in protein folding, *Nat Struct Biol 6*, 554-556.

Chapter 4

Cooperative Cold Denaturation: The Case of the C-terminal Domain of the Ribosomal Protein L9

Abstract

Cold denaturation is a general property of globular proteins, but it is difficult to directly characterize since the transition temperature of protein cold denaturation, T_c, is often below the freezing point of water. As a result, studies of protein cold denaturation are often facilitated by addition of denaturants, or by using destabilizing pHs, or extremes of pressure, or by reverse micelle encapsulation, and there are few studies of cold induced unfolding under near native conditions. The thermal and denaturant induced unfolding of single domain proteins is usually cooperative, but the cooperativity of cold denaturation is controversial. The issue is of both fundamental and practical importance since cold unfolding may reveal information about otherwise inaccessible partially unfolded states and because many therapeutic proteins need to be stabilized against cold unfolding. It is thus desirable to obtain more information about the process under non-perturbing conditions. The ability to access cold denaturation in native buffer is also very useful for characterizing protein thermodynamics, especially when other methods are not applicable. In this work, we study a point mutant of the C-terminal domain of the ribosomal protein L9 (CTL9) which has a T_c above 0 °C. The mutant was designed to enable the study of cold denaturation under near native conditions. The cold denaturation process of I98A CTL9 was characterized by NMR, CD, and FTIR. The results are

consistent with apparently cooperative, two-state cold unfolding. SAXS studies show that the unfolded state expands as the temperature is lowered.

Note: The studies presented in this chapter have been published (Luan, B., Shan, B., Baiz, C., Tokmakoff, A, and Raleigh D. P. (2013) Cooperative Cold Denaturation: The Case of the C-terminal Domain of the Ribosomal Protein L9. Biochemistry, 52: 2402-2409). This chapter contains direct excerpts from the paper with suggestions and revisions by Prof. Daniel P. Raleigh. The FTIR experiment was performed by Dr. Carlos Baiz and Prof. Andrei Tokmakoff. I thank Dr. Vadim Patsalo for helpful discussions on the SVD analysis of the far-UV CD data and Dr. Marc Allaire for the help of SAXS experiments.

4.1. Introduction

Protein cold denaturation, a transition from the folded state to an unfolded state induced by lowering the temperature from the temperature of the maximum stability, is a general property of globular proteins. This phenomenon is well predicted by the Gibbs-Helmholtz equation, and is rationalized by the decrease of hydrophobic interactions together with solvation effects including water penetration, although the details are still under debate.[1-10] There is evidence that cold denaturation is relevant *in vivo* and cold denaturation has important practical implications for the formulation of biotheraputics.[11-16] For example, cold unfolding has been reported to affect monoclonal antibodies.[11-13] Proteins undergo ice-water surface denaturation, cold denaturation, and cryoconcentration, making it critical to carefully design freeze-thaw and storage conditions in order to maintain the activity of protein pharmaceuticals.[16] Studies of protein function in cold-adapted organisms have revealed a delicate balance between harsh environments (cold-induced denaturation) and catalytic activity (the compromise of structural flexibility).[14, 15] Thus a more detailed understanding of the process is of considerable practical importance.

Cold denaturation of amyloid has been reported,[17] indicating it is a general phenomenon. Interestingly, intrinsically disordered proteins (IDPs) have been suggested to be more resistant to cold-treatment,[18] although this will likely depend on the particular protein under investigation. Cold denaturation has been proposed to provide access to important partially unfolded states that are otherwise inaccessible.[19-21] Analysis of the cold

unfolding transition also allows one to characterize protein stability, especially for proteins whose thermal unfolding transitions are hard to measure.

Unfortunately, the transition temperature of cold denaturation is often well below the freezing point of aqueous solution, making it difficult to study, and this has limited progress. Modifications to the system, such as adding denaturant, high-pressure, extremes of pH,[9, 22-28] or encapsulating the protein of interest inside micelles,[20, 29, 30] have enabled studies of the cold denatured state, but those conditions are different from native buffer, and proteins may behave differently which subjected to strongly non-native conditions.

Two-state, cooperative thermal and denaturant induced unfolding is a common feature for small single domain proteins, however, the cooperativity of cold denaturation is less well characterized and is controversial.[19-21, 31-33] In order to obtain more insight into the cooperativity of protein cold-denaturation process it is desirable to study systems under near physiological conditions. A very limited number of such studies have been reported.[34-37] Here we characterize the cold unfolding of a globular protein in native buffer and show that it is an apparent two-state, cooperative process. A point mutant of the C-terminal domain of the ribosomal protein L9 (CTL9) is used as a model protein in this work. The 92-residue domain adopts an α-β fold (**Figure 4-1**) and its thermal, pH induced, and denaturant induced unfolding have been well characterized, and appear to be two-state.[38-40] The T_m is pH dependent, owing in part to several buried histidine side-chains. The conformational properties of the cold unfolded state of the mutant have been

examined at 12 °C,[41] but the cooperativity of the transition has not been probed nor have the properties of the cold unfolded state been analyzed in detail at lower temperatures.

4.2. Materials and methods

4.2.1. Mutagenesis, protein expression and purification

I98A CTL9 was expressed and purified using procedures previously described for wild type CTL9.[38, 42] The identity of the protein was confirmed by DNA sequencing and MALDI-TOF mass spectroscopy. The observed molecular weight was 9940.9 ± 1.2 Da and the expected molecular weight was 9939.5 Da. The yield of the protein was 70-80 mg/L in LB media and 30 mg/L in M9 minimal media. A similar yield was observed for wild type CTL9. The purity was tested by analytical HPLC.

4.2.2. Nuclear magnetic resonance (NMR) experiments

^{15}N-labeled I98A CTL9 was dissolved in 10 % D_2O and 90 % H_2O with 10 mM MOPS and 150 mM NaCl at a protein concentration of around 1.0 mM. The pH was adjusted to 6.0. ^{15}N-^{1}H correlated hertonuclear single coherence (HSQC) experiments were performed on a 600 MHz Varian spectrometer, from 5 ˚C to 25 ˚C with an increment of 3 ˚C to 4 ˚C. The temperature was calibrated using a standard methanol sample. ^{15}N-^{1}H HSQC spectra were recorded using 2048 x 512 complex points with 16 scans per increment and spectral widths of 8,000.0 Hz and 2,200.0 Hz for the ^{15}N and ^{1}H dimensions, respectively. The ^{15}N offset frequency was set to 118.0 ppm and the ^{1}H

dimension was centered at the water resonance. The spectra were processed using NMRpipe software[43] and visualized in NMRViewJ.[44] 1D NMR experiments were performed on a sample containing 1.0 mM I98A CTL9 in 10 mM MOPS and 150 mM NaCl, dissolved in 100 % D_2O with the pD adjusted to 5.6 (uncorrected pH meter reading). 0.5 mM 2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) was used as an internal reference. The data was analyzed using the software package Mnova 7.

4.2.3. Circular dichroism (CD) spectroscopy

CD experiments were performed on a Chirascan CD spectrometer. The protein was dissolved in 10 mM MOPS and 150 mM NaCl buffer in D_2O, at a protein concentration of about 20 μM. D_2O was used to allow comparison with the ^1H NMR spectra and FTIR experiment. Far-UV wavelength spectra were collected in a 1 mm cuvette from 196 nm to 260 nm with a 1 nm increment and averaged with 3 repetitions. Thermal denaturation experiments were conducted as a function of pD between 4.0 and 8.0. Each thermal denaturation was performed by monitoring the ellipticity at 222 nm in a 1 cm cuvette, from 4 ˚C to 98 ˚C, with a 2 ˚C step and a heating rate of 1 ˚C/min. SVD analysis was carried out using the program implemented in the R software version 2.13.0 (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). The pD 4.0 thermal denaturation data was fit to a quadratic equation to obtain the unfolded state signal. The pD 8.0 curve was fit to the following equation to obtain thermodynamic parameters of the native state. The observed CD signal, $\theta(T)$, was fit to the equation:

$$\theta(T) = \frac{(a_n + b_n T) + (a_d + b_d T)\exp(-\Delta G_u^\circ(T)/RT)}{1 + \exp(-\Delta G_u^\circ(T)/RT)} \quad \text{(eq 4-1)}$$

where $\Delta G_u^\circ(T)$ is the free energy change upon thermal unfolding described by the Gibbs-Helmholtz equation:

$$\Delta G_u^\circ(T) = \Delta H^\circ(T_m) - T\Delta S^\circ(T_m) + \Delta C^\circ_P [T - T_m - T \ln(T/T_m)] \quad \text{(eq 4-2)}$$

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 function of temperature. T_m is the thermally induced unfolding midpoint temperature, $\Delta H^\circ(T_m)$ is the enthalpy change at T_m. The signal expected for a fraction of folded of 0.5 was estimated by taking the average of the native state baseline and the unfolded state baseline as a function of temperature.

4.2.4. FTIR experiments

Amide I IR spectra were collected over the temperature range of 3 ˚C to 25 ˚C at 1 ˚C intervals, using a Nicolet 380 FTIR with a resolution of 2 cm^{-1}. The sample cell consists of CaF_2 windows separated by a 50 μm PTFE spacer. 10mM MOPS buffer with 100 mM NaCl in 100 % D_2O at pD 5.6 was used for the FTIR measurements. The concentration of I98A CTL9 was 5.0 mg/mL. To account for the thermal shift of the amide I band, which is independent of structural changes in the protein, FTIR spectra were shifted by a phenomenological value of 0.052 cm^{-1}/°C with a reference temperature of 15 °C.

Difference spectra were obtained by subtracting the final spectrum collected at 25 ˚C in order to obtain the difference between the native and DSE FTIR contributions. Singular value decomposition was performed on the equilibrium spectra in order to project out the spectral response, and temperature profile associated with the cold denaturation of CTL9. The FTIR data were collected by Dr. Carlos Baiz from Prof. Andrei Tokmakoff group, and the data is included in this chapter for the purpose of the completeness of presenting this project.

4.2.5. Small angle X-ray scattering measurements

Samples of I98A CTL9 were prepared in buffer consisting 10 mM MOPS and 150 mM NaCl in 100 % H_2O, with the pH adjusted to 6.0. Scattering experiments were performed at beamline X9 at Brookhaven National Laboratory, National Synchrotron Light Source I (Upton, New York, USA). Protein samples were injected into a 1 mm capillary continuously during the measurement at a rate of 0.67 μL/s in order to avoid radiation damage. The exposure time for each measurement was 30 s. Scattering data was collected for I98A CTL9 at a protein concentration of 3.75 mg/mL, at 7 °C, 12 °C, and 25 °C. Each sample was measured three times and then averaged before data analysis. The program pyXS (http://www.bnl.gov/ps/x9/software/pyXS.asp) was used for buffer subtraction, and the radius of gyration (R_g) was obtained using the Guinier approximation using the program PRIMUS,[45]

$$I(q) = I(0) \bullet \exp(-R_g^2 q^{2/3}) \qquad \text{(eq 4-3)}$$

where $I(q)$ is the intensity at scattering vector q.[46]

4.3. Results

4.3.1. I98A CTL9 undergoes cold denaturation under near native conditions, with a T_c above 0 °C

We used a designed core mutant of CTL9, chosen to destabilize the protein without perturbing its fold, or the two-state nature of the high temperature unfolding transition [41]. I98 lies in the hydrophobic core of the protein (**Figure 4-1**), and its truncation to a smaller hydrophobic residue destabilizes the domain by reducing the hydrophobic driving force for folding and by altering core packing. The I98A mutant destabilizes CTL9 by about 4 kcal/mol. The temperature of cold denaturation increases as ΔC_p^o increases and as $\Delta H^o(T_m)$, the enthalpy change at the midpoint of thermal unfolding, decreases. The I98A mutant was chosen because the altered core packing was expected to decrease $\Delta H^o(T_m)$.[40]

The mutation was also designed to increase ΔC_p^o by potentially weakening hydrophobic clusters in the unfolded state.[40] The structure of the mutant is the same as wild type, as suggested by CD and NMR chemical shift analysis. The classical approach to probe the cooperativity of folding or unfolding is to use two or more distinct, structurally sensitive, spectroscopic methods. Different structural probes will yield overlapping transition curves if only two distinct structural states are well populated during the protein folding

process.[47, 48] In the present work, we use NMR, CD, and FTIR to follow the unfolding of the I98A mutant.

^{15}N-^{1}H HSQC spectra were collected for the CTL9 I98A mutant at 5 ˚C and 25 ˚C at pH 6.0 (**Figure 4-2**). The resonances due to the unfolded state have limited dispersion in both the ^{15}N and ^{1}H dimensions, which is typical of an unfolded protein ensemble. At low temperatures the peaks from the denatured state ensemble (DSE) are much more intense than those from the folded state, indicating that the DSE is the dominant species under these conditions. Peaks from the folded state can be observed even in the 5 ˚C spectrum at lower contour levels. The line-widths are sharp, suggesting that the protein is monomeric, in agreement with hydrodynamic measurements.[40] The peaks of the DSE at 25 °C match well with those of the cold denatured state at low temperature. Most of the native resonances in the 25°C HSQC spectrum of the mutant are not shifted relative to their position in the spectrum of the folded wild type protein,[41] indicating that the mutation does not significantly perturb the structure. No obvious broadening of the resonances is detected for any of the peaks, demonstrating that the two states are in slow exchange on the NMR chemical shift time scale.

Like wild type CTL9, the stability of the I98A mutant depends strongly on pH. The protein becomes less stable when the histidine side-chains are protonated at lower pH and is more stable when the histidine side-chains are deprotonated. **Figure 4-3** shows thermal unfolding curves for I98A CTL9 at different pD values detected by CD. D_2O is used for these studies to allow direct comparisons with FTIR and ^{1}H NMR. Cold denaturation is

observed between pD values of 4.7 and 6.6. The shape of the curves is similar to those observed in H_2O,[41] although the transition temperatures are shifted. The red curve represents a quadratic fit to the pD 4.0 data and provides an experimental estimation of the DSE signal as a function of temperature, since the fraction folded is zero under these conditions. The pD 8.0 data provides an estimation of the CD signal for the fully folded state as the fraction folded is one at 25 ˚C at pD 8.0. The solid black line represents an extrapolation of the folded baseline. The green curve represents the signal expected for a fraction folded of 0.5, i.e. the midpoint of the transition. Based on this dataset, the fraction folded (or unfolded) can be easily calculated at given temperature and pD. At pD 5.6 the population of the cold denatured ensemble is 76 % at 4˚C, and at 25˚C the native state and the DSE are populated to 64 % and 36 %, respectively.

4.3.2. CD, FTIR, and NMR are consistent with two-state, cooperative cold denaturation

Far-UV CD spectra of I98A CTL9 were recorded over the temperature range of 2 ˚C to 25 ˚C (**Figure 4-4**). Data were collected in D_2O to allow comparison with FTIR and 1D ^{1}H-NMR measurements. D_2O can stabilize proteins so it is important to match the isotopic composition of the solvent.[49, 50] The spectrum at 25 ˚C indicates a mixture of α-helix, β-strand, and coil. At temperatures below 10 ˚C, the spectra are typical of those expected for an unfolded protein, but do not correspond to a classic random coil. In particular, there is still significant intensity at 222 nm, consistent with the presence of residual helical structure. An isodichroic point at 207 nm is observed during cold

unfolding, consistent with a two-state transition. An isodichroic point is a necessary, but not a sufficient condition for a two-state transition. Single value decomposition (SVD) analysis shows that only two major spectral components are needed to define the transition: the second component is weighted 12 % relative to the largest component, and the third contributes 1.1 %, all of the other components are negligible.

FTIR data were collected for the CTL9 I98A mutant over the temperature range of 3 ˚C to 25 ˚C in D_2O (**Figure 4-5**). An isosbestic point is observed and SVD analysis of the FTIR data shows that only two significant components are required to describe the difference spectra. The second and the third components weights are 0.92 % and 0.28 % of the first component respectively. Difference spectra show a loss feature in around 1618 cm^{-1}, typically associated with proline turns and β-structure, as well as a gain in the 1630 cm^{-1} region, associated with the strong ν perpendicular mode of anti-parallel β-sheets.[51, 52] A pronounced loss of intensity around 1650 cm^{-1} can be attributed to modes arising from α-helices and disordered regions. The spectra indicate that the DSE contains β-sheet character, and decreased α-helix content in comparison to the native structure.

The aromatic region of the ^1H NMR spectra of I98A CTL9 displays well resolved peaks from tyrosine resonances of the native state and the DSE (**Figure 4-6**). The signal tyrosine is at residue-126 which is located in the second α-helix. Integration of the area under the unfolded and folded tyrosine peaks provides an independent estimate of the fraction unfolded as a function of temperature. The values are in good agreement with the ones obtained by CD, consistent with a two-state process (**Figure 4-7**).

As a control we compared pD induced unfolding at 25 °C to thermal unfolding experiments at different pD values. This produces a test of the reliability of the parameters extrapolated from the thermal unfolding data. Each thermal unfolding curve was analyzed using the Gibbs-Helmholtz equation to obtain the thermodynamic parameters at 25 °C. There is excellent agreement between these values and the values obtained by the pD titration unfolding curve (**Figure 4-8**). The strong agreement indicates that the population estimates obtained from extrapolation of the thermal unfolding curves are precise. The agreement is also consistent with two-state unfolding.

4.3.3. Small angle X-ray scattering data show that the cold denatured state expands at low temperature

To further characterize the cold denatured state of I98A CTL9, we collected SAXS data at 7, 12, and 25 °C. The population of the native state at those temperatures is 36, 50, and 65 %, respectively. We used two methods to estimate R_g of the DSE of I98A CTL9. First we used the experimental curve and subtracted the scattering profile for the native state. By subtracting the signal of the native state, the R_g of unfolded I98A CTL9 can be estimated. The folded state curve was collected independently for the fully folded state of wild type CTL9 (**Figure 4-9**), and was subtracted, with appropriate weighting, from the experimental curve of I98A CTL9 (**Figure 4-10**). The R_g values of the cold denatured

I98A-CTL9 were obtained by the Guinier analysis (**Figure 4-11**) The second method estimated R_g for the DSE using the standard relationship for a two-component system:

$$R_g^2{}_{observed} = p_{native} \cdot R_g^2{}_{native} + p_{DSE} \cdot R_g^2{}_{DSE} \quad (1)$$

p_{native} is the population of the native state and p_{DSE} is the population of the DSE ensemble.[53] Using the known R_g value for the native state provided by wild type CTL9, and the relative populations of the native and DSE allows R_g of the DSE to be estimated. The two approaches give estimates which are in very good agreement (**Table 4-1.**). The cold denatured state expands, as judged by R_g, as the temperature is lowered, increasing by 17 % to 27 % as the temperature is reduced from 25 ºC to 7 °C. The increase is significant and larger than the estimated uncertainty. This confirms previous hydrodynamic studies of I98A CTL9.[40] Control experiments show that the values of R_g for the folded state is independent of the temperature.

4.4. Discussion

The denaturant induced and thermally induced unfolding of globular single domain proteins is usually cooperative, but the cooperativity of cold denaturation is less certain. The Yfh1 protein, a small α-β protein, undergoes two-state cold unfolding, however, deviations from two-state cold unfolding have been reported for ubiquitin encapsulated in reverse micelles.[19-21, 34-36] Temperature dependent NMR, CD and FTIR experiments, together with the SVD analysis, are all consistent with cooperative cold denaturation of I98A CTL9 in native buffer.

Why do the results reported here differ from studies of ubiquitin encapsulated in reverse micelles? There may be fundamental differences in the behavior of the two proteins, although the equilibrium thermal unfolding of both have been reported to be cooperative in homogeneous solution. It has been suggested that studies in reverse micelles can be complicated by temperature dependent interactions between the protein and the micelle or by water shedding.[32, 54] PFG-NMR diffusion experiments[40] and the SAXS data reported here demonstrate that the cold denatured states of proteins can expand at low temperatures, which may enhance the opportunity for interactions between the micelle and the protein. Halle and coworkers have monitored the hydration dynamics of ubiquitin in non-perturbing picoliter emulsion droplets using water-^{17}O spin relaxation.[5] Ubiquitin was found to be thermodynamically stable even at -32 ˚C, suggesting that the cold denaturation of ubiquitin encapsulated in reverse micelles might be induced by the low water content in the micelles rather than by low temperature. Irrespective of the details of previous studies, the present work produces a system in which cold unfolding can be observed in native buffer.

The cooperativity of cold unfolding is important from a basic protein thermodynamics perspective, but it also has practical implications. Protein cold denaturation is an issue in the food processing industry, in cryopreservation and in protein pharmaceuticals. Protein based therapeutics are usually stored at low temperature and there have been reports of the cold denaturation of monoclonal antibodies.[11-13] Thus, designing resistance to cold induced unfolding is of practical interest. It is conceptually easier to stabilize a

cooperatively folded system rather than ones which folds non-cooperatively. In the former the effects of a mutation contribute to the global stability, while in the latter substitutions may impact the local stability of small portions of the structure, but not affect other regions. It is clearly more challenging to design mutations that stabilize the native ensemble for the latter class of protein.

The CTL9 mutant analyzed here and Yfh1, in low salt, appear to undergo cooperative cold unfolding in native buffer in the absence of denaturation.[37] Whether cooperative cold denaturation is a general property of globular proteins remains to be determined.

Table 4-1. R_g values for I98A CTL9.

Temperature (˚C)	R_g, folded (Å) [1)]	R_g, unfolded (Å) [2)]	R_g, unfolded (Å) [3)]
7	15.2 ± 0.3	28.4 ± 0.9	28.2 ± 1.4
12	14.8 ± 0.4	26.1 ± 1.1	25.1 ± 1.6
25	14.5 ± 0.3	24.0 ± 1.1	22.2 ± 1.7

1) Measured using wild type CTL9.
2) Calculated from the observed scattering curve after subtraction of the folded state scattering curve.
3) Calculated from the equation: $R_g^2{}_{observed} = p_{native} \cdot R_g^2{}_{native} + p_{DSE} \cdot R_g^2{}_{DSE}$, where p_{native} and p_{DSE} are the fractional populations of the native state and the DSE.

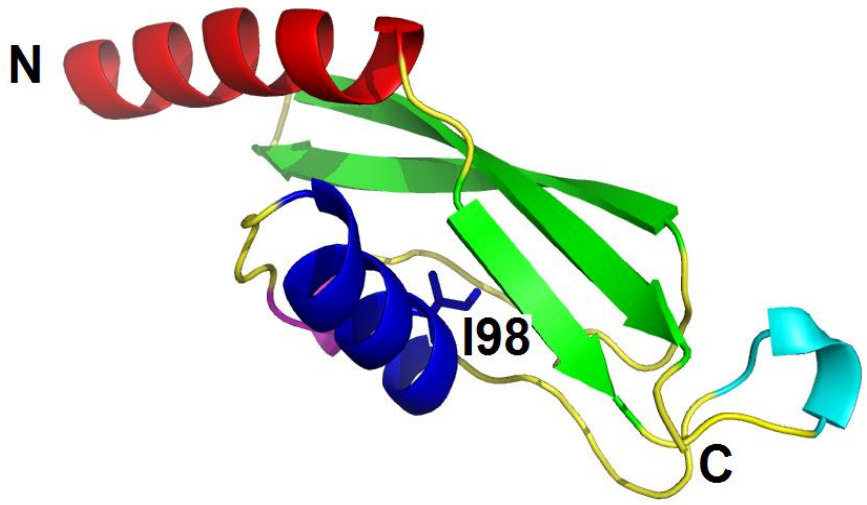


Figure 4-1. Ribbon diagram of CTL9. The β-sheet is green, the first α-helix red and the second α-helix blue. The hydrophobic core residue I98 and the two termini are labeled. The PDB file entry is 1DIV. The diagram was constructed using program PyMol.

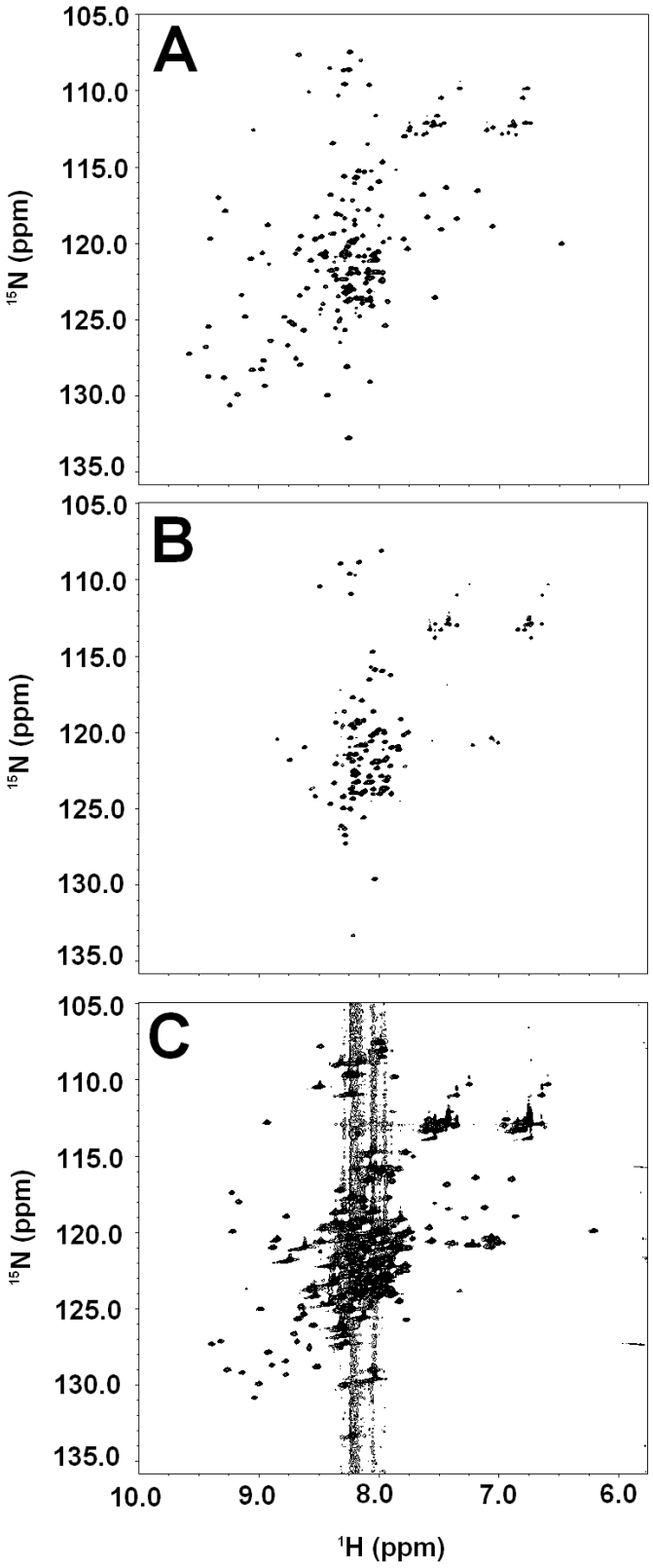

A
B
C
105.0
110.0
115.0
120.0
125.0
130.0
135.0
^{15}N (ppm)
10.0
9.0
8.0
7.0
6.0
^{1}H (ppm)

Figure 4-2. ^{15}N-^{1}H HSQC spectra of I98A CTL9 at pH 6.0, in 10 mM MOPS and 150 mM NaCl buffer, 10 % D_2O and 90 % H_2O. (A) 25 °C, (B) 5 °C, (C) The 5 °C spectra plotted at lower contour level in order to visualize the folded state resonances.

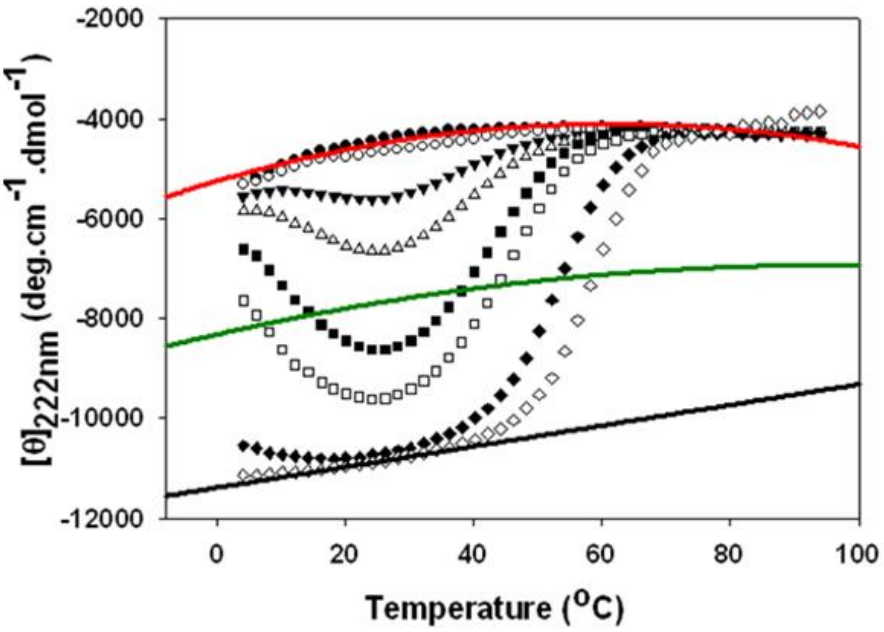


Figure 4-3. CD detected thermal unfolding curves for I98A CTL9 monitored at different pD values for protein dissolved in 10 mM MOPS and 150 mM in D_2O. (●) pD 4.0, (○) pD 4.7, (▼) pD 5.2, (△) pD 5.4, (■) pD 5.6, (□) pD 6.0, (◆) pD 6.6, and (◇) pD 8.0. Note the strong dependence of the transitions on pD. The red curve represents a quadratic fit to the pD 4.0 data and provides an experimental estimation of the unfolded state signal as a function of temperature since the fraction folded is zero under these conditions. The pD 8.0 data provides an estimation of the CD signal for the fully folded state as the fraction folded is 1.0 at 25 °C, pD 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 0.5, i.e. the midpoint of the transition.

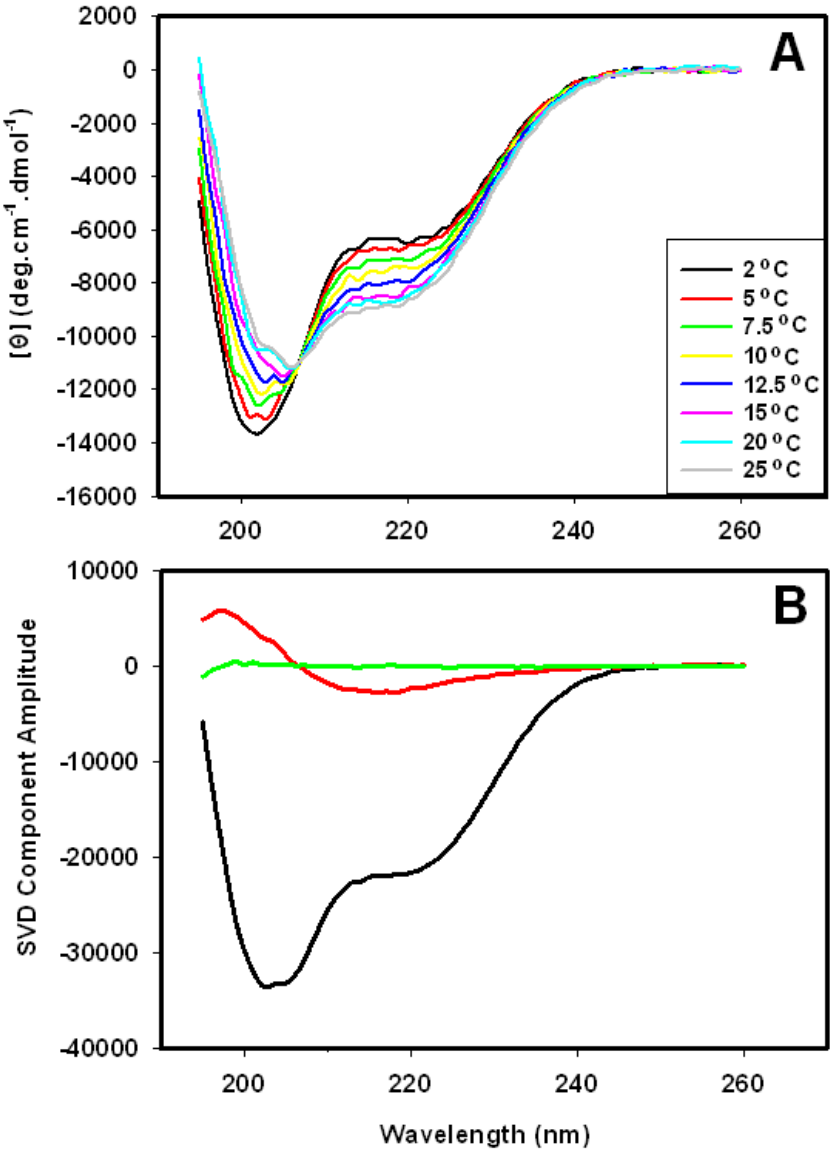


Figure 4-4. (A) Far-UV CD spectra of the CTL9 I98A mutant as a function of temperature, recorded over the temperature range from 2 °C to 25 °C at pD 5.6, in 10 mM MOPS and 150 mM NaCl. (B) Results of the SVD analysis for the temperature dependent unfolding of I98A CTL9 monitored by CD. The relative amplitudes of the first (black), second (red), and the third (green) components are 1.00, 0.120, and 0.011, respectively.

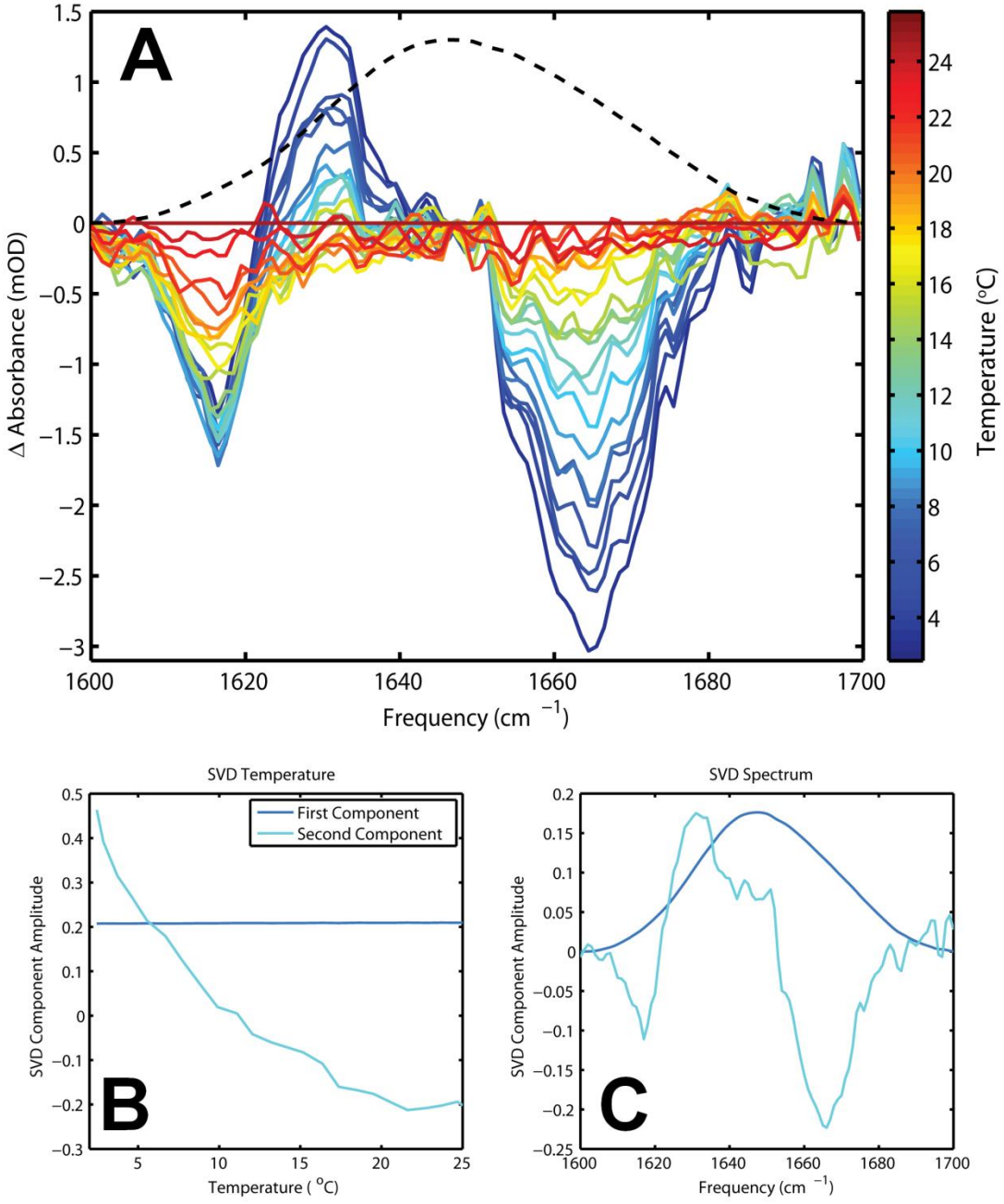

A
Δ Absorbance (mOD)
Frequency (cm −1)
Temperature (°C)
B
SVD Temperature
First Component
Second Component
SVD Component Amplitude
Temperature (°C)
C
SVD Spectrum
SVD Component Amplitude
Frequency (cm −1)

Figure 4-5. (A) Difference FTIR spectra of CTL9 I98A at pD 5.6 from 3 ˚C to 25 ˚C. The 25 ˚C data was subtracted from the data collected at different temperatures. (B) Temperature dependence of the SVD components suggests that the first component is constant and the second component shows monotonic decrease, which is consistent with cooperative cold unfolding. (C) SVD analysis of the FTIR data shows two major components for the cold denaturation of CTL9 I98A. The relative amplitudes of the first (blue), second (cyan) components are 1.00, 0.0921 respectively. The relative amplitude of the third component (not shown) is 0.0282.

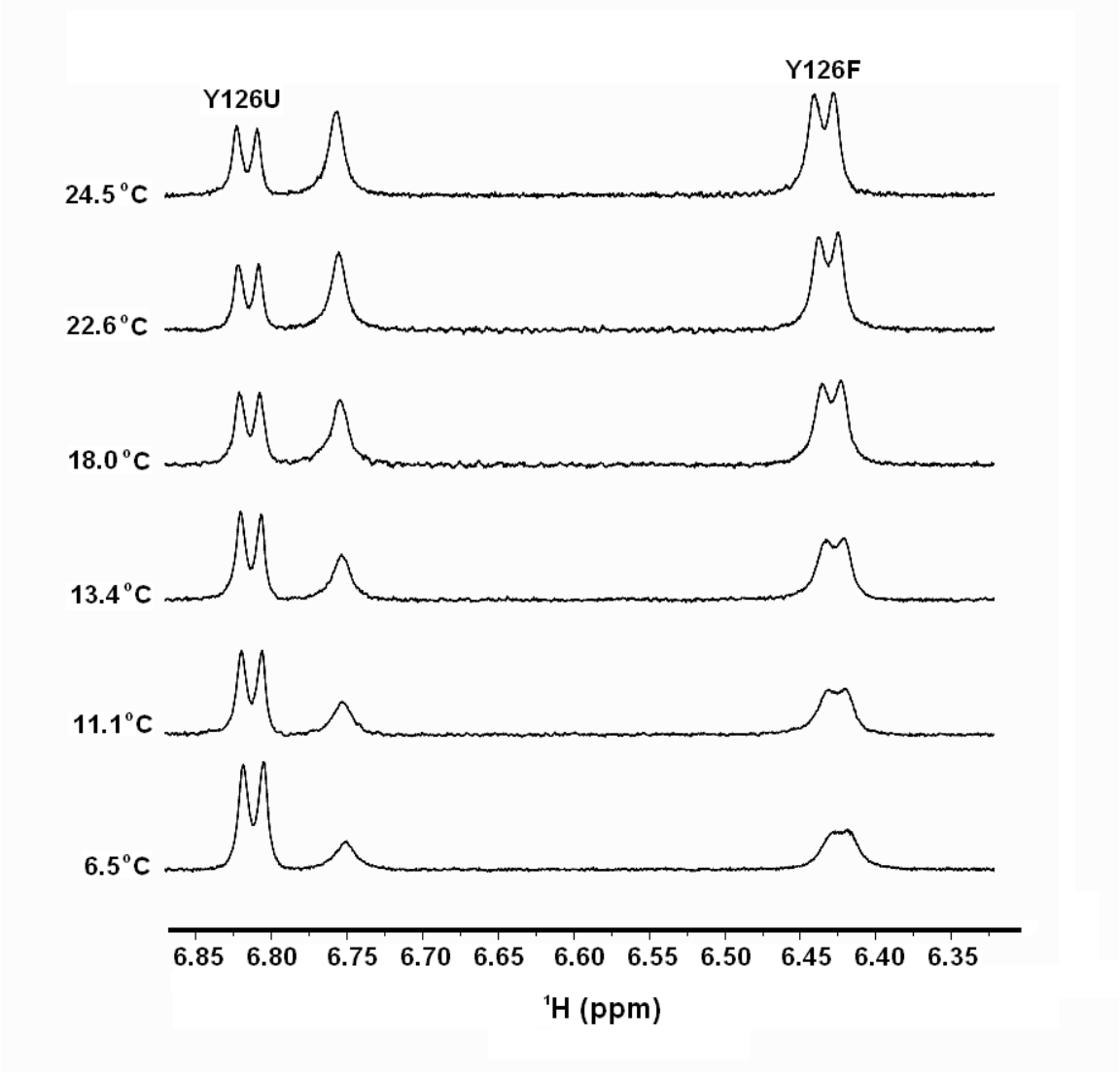


Figure 4-6. 1D ^{1}H-NMR spectra of I98A CTL9 at different temperatures in 10 mM MOPS and 150 mM NaCl, at pD 5.6, in 100 % D_2O. Only the aromatic region is shown and the resonance assignments are labeled.

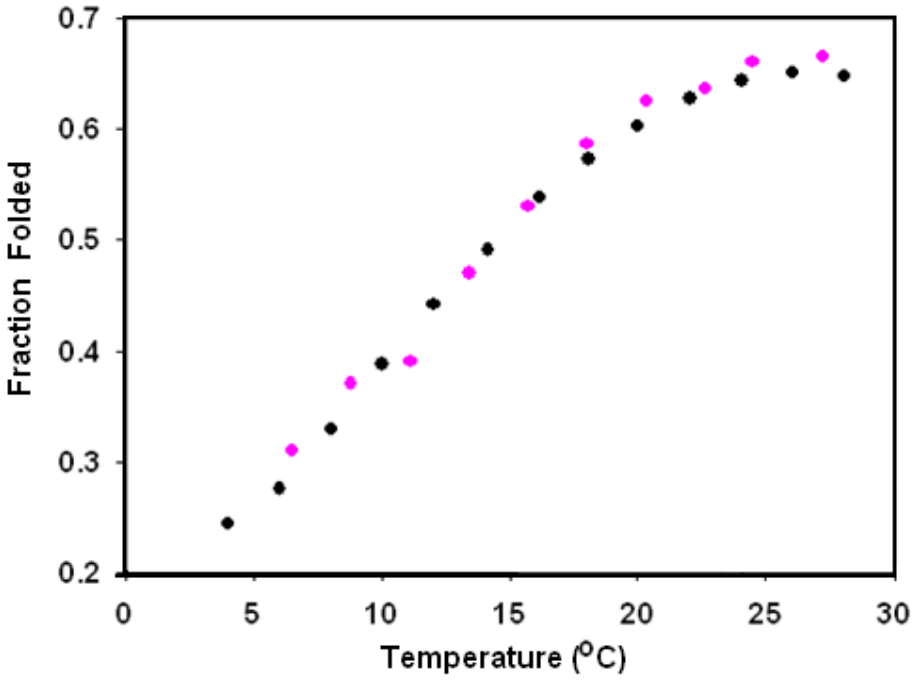


Figure 4-7. Estimation of the fraction of folded versus temperature based on CD data (•) calculated from the CD temperature melting curve at pD 5.6 and NMR data (•) calculated from the relative peak intensities of the Y126 resonance from unfolded and folded states. Experiments were performed as a function of temperatures in 10 mM MOPS and 150 mM NaCl, at pD 5.6, in 100 % D_2O.

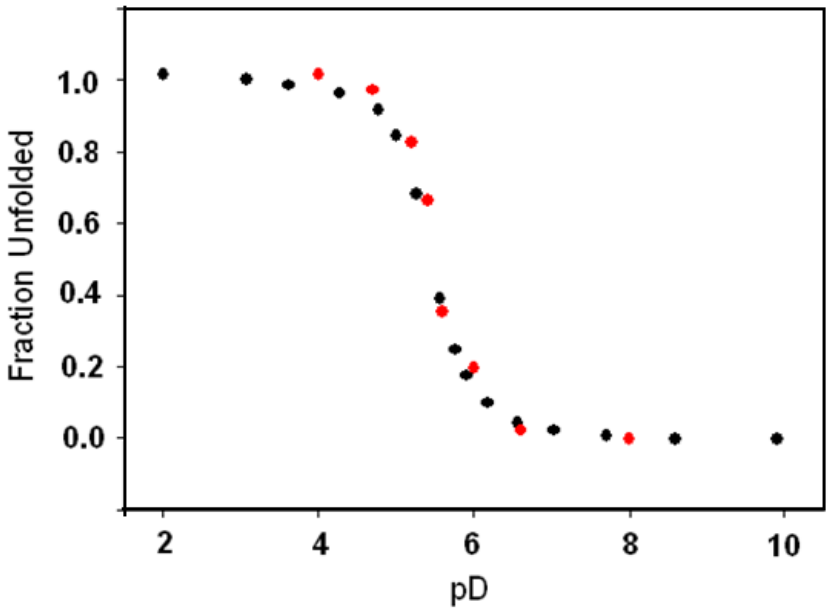


Figure 4-8. pD-induced denaturation of I98A CTL9. The CD signal at 222 nm was monitored in 10 mM MOPS and 150 mM NaCl in D_2O. The pD values were adjusted by adding DCl and NaOD directly into the cuvette. (•) calculated from the pD-titration curve, (•) calculated from the extrapolation of thermal unfolding curves collected as a function of pD. The good agreement validates the extrapolation of the thermal unfolding data and suggests that the population estimates are precise, and that the unfolding is two-state.

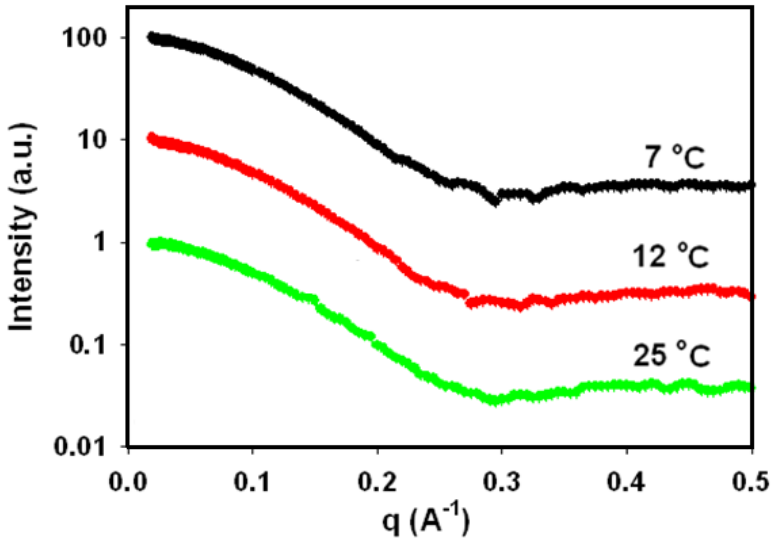


Figure 4-9. Scattering curves from wild type CTL9 (control experiments) at different temperatures: 7 °C (●), 12 °C (●), and 25 °C (●). No apparent change in the shape of the curve was observed as a function of temperature, and the calculated R_g values are in excellent agreement, ranging from 14.9 ± 0.3 Å to 15.2 ± 0.3 Å. The curves are offset from each other for clarity. The buffer was 10 mM MOPS and 150 mM NaCl, in H_2O.

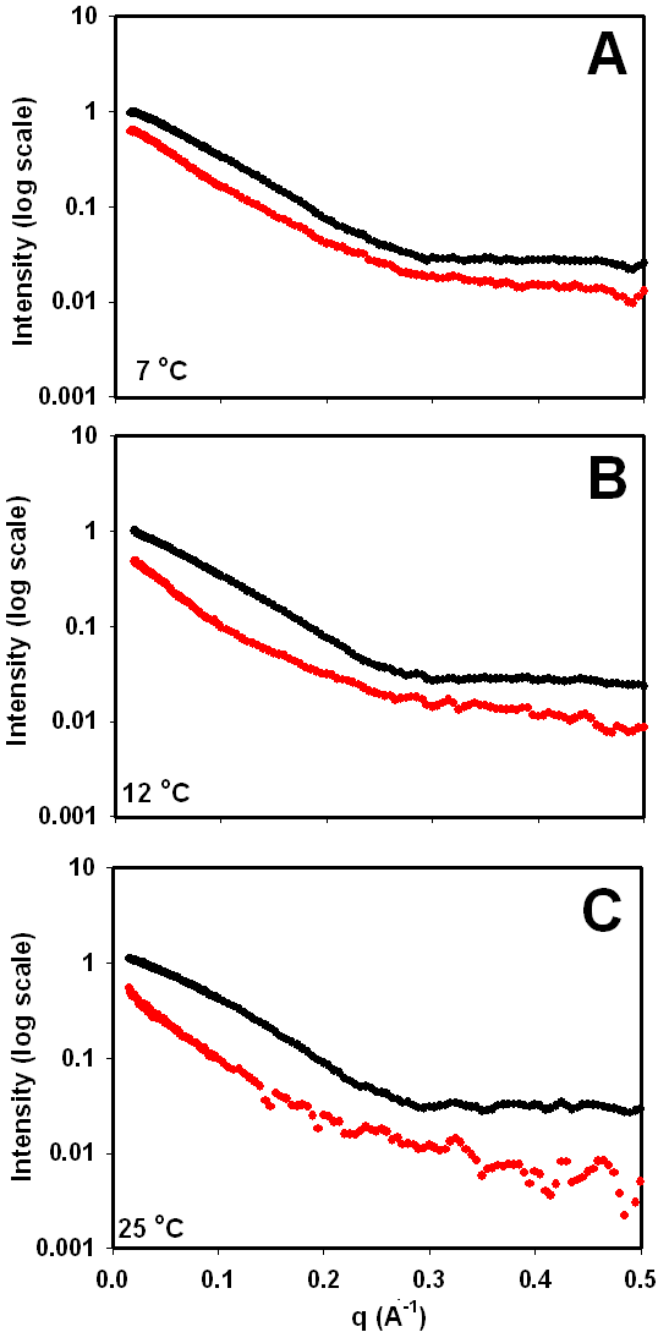


Figure 4-10. The observed scattering curves for CTL9 I98A (●) and the profile after subtraction (●) of the contribution of the folded ensemble. (A) 7 °C; (B) 12 °C; (C) 25 °C.

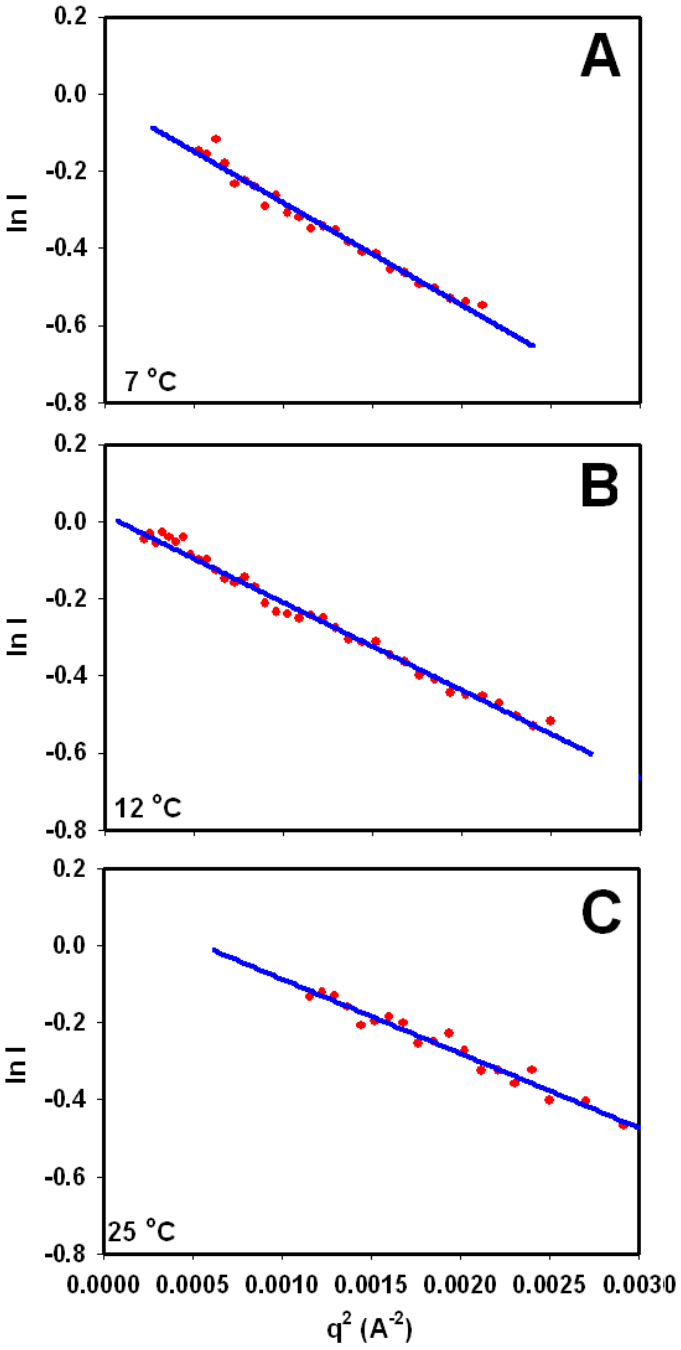


Figure 4-11. Guinier analysis of the profiles shown in Figure 9. Scattering profiles for the cold denatured ensembles (•) are shown. The Guinier approximation is shown as a blue straight line. (A) 7 °C; (B) 12 °C; (C) 25 °C.

4.5. References:

[1] Graziano, G. (2010) On the molecular origin of cold denaturation of globular proteins, *Phys Chem Chem Phys 12*, 14245-14252.

[2] Yoshidome, T., and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein, *Phys Rev E 79*, 030905.

[3] Dias, C. L. (2012) Unifying microscopic mechanism for pressure and cold denaturations of proteins, *Phys Rev Lett 109*, 048104.

[4] Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I., and Grant, M. (2008) Microscopic mechanism for cold denaturation, *Phys Rev Lett 100*, 118101.

[5] 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.

[6] Dias, C. L., Ala-Nissila, T., Wong-ekkabut, J., Vattulainen, I., Grant, M., and Karttunen, M. (2010) The hydrophobic effect and its role in cold denaturation, *Cryobiology 60*, 91-99.

[7] Tsai, C. J., Maizel, J. V., Jr., and Nussinov, R. (2002) The hydrophobic effect: A new insight from cold denaturation and a two-state water structure, *Crit Rev Biochem Mol Biol 37*, 55-69.

[8] Franks, F. (1995) Protein destabilization at low temperatures, *Adv Protein Chem 46*, 105-139.

[9] Privalov, P. L., Griko Yu, V., Venyaminov, S., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin, *J Mol Biol 190*, 487-498.

[10] Lopez, C. F., Darst, R. K., and Rossky, P. J. (2008) Mechanistic elements of protein cold denaturation, *J Phys Chem B 112*, 5961-5967.

[11] Todgham, A. E., Hoaglund, E. A., and Hofmann, G. E. (2007) Is cold the new hot? Elevated ubiquitin-conjugated protein levels in tissues of Antarctic fish as evidence for cold-denaturation of proteins *in vivo*, *J Comp Physiol B 177*, 857-866.

[12] Lazar, K. L., Patapoff, T. W., and Sharma, V. K. (2010) Cold denaturation of monoclonal antibodies, *mAbs 2*, 42-52.

[13] Feller, G. (2010) Protein stability and enzyme activity at extreme biological temperatures, *J Phys Condens Mat 22*, 323101-323117.

[14] Pucciarelli, S., Parker, S. K., Detrich, H. W., 3rd, and Melki, R. (2006) Characterization of the cytoplasmic chaperonin containing TCP-1 from the Antarctic fish *Notothenia coriiceps, Extremophiles 10*, 537-549.

[15] Hayley, M., Chevaldina, T., and Heeley, D. H. (2011) Cold adaptation of tropomyosin, *Biochemistry 50*, 6559-6566.

[16] Kolhe, P., and Badkar, A. (2011) Protein and solute distribution in drug substance containers during frozen storage and post-thawing: A tool to understand and define freezing-thawing parameters in biotechnology process development, *Biotechnol Progr 27*, 494-504.

[17] Kim, H. Y., Cho, M. K., Riedel, D., Fernandez, C. O., and Zweckstetter, M. (2008) Dissociation of amyloid fibrils of α-synuclein in supercooled water, *Angew Chem Int Ed Engl 47*, 5046-5048.

[18] Tantos, A., Friedrich, P., and Tompa, P. (2009) Cold stability of intrinsically disordered proteins, *FEBS Lett 583*, 465-469.

[19] 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.

[20] 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.

[21] 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.

[22] Tamura, A., Kimura, K., Takahara, H., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .1. CD and DSC Studies, *Biochemistry 30*, 11307-11313.

[23] Tamura, A., Kimura, K., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .2. NMR Studies, *Biochemistry 30*, 11313-11320.

[24] Chen, B. L., and Schellman, J. A. (1989) Low-temperature unfolding of a mutant of phage-T4 lysozyme .1. Equilibrium studies, *Biochemistry 28*, 685-691.

[25] Chen, B. L., Baase, W. A., and Schellman, J. A. (1989) Low-Temperature Unfolding of a Mutant of Phage-T4 Lysozyme .2. Kinetic Investigations, *Biochemistry 28*, 691-699.

[26] Griko, Y. V., Privalov, P. L., Sturtevant, J. M., and Venyaminov, S. Y. (1988) Cold denaturation of staphylococcal nuclease, *Proc Natl Acad Sci USA 85*, 3343-3347.

[27] Pace, N. C., and Tanford, C. (1968) Thermodynamics of unfolding of β-lactoglobulin a in aqueous urea solutions between 5 and 55 degrees, *Biochemistry 7*, 198-206.

[28] Jacobsen, C. F., and Christensen, L. K. (1948) Influence of temperature on the urea denaturation of β-lactoglobulin, *Nature 161*, 30-31.

[29] 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.

[30] 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.

[31] Tian, J. H., and Garcia, A. E. (2011) Simulations of the confinement of ubiquitin in self-assembled reverse micelles, *J Chem Phys 134*, 225101.

[32] 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.

[33] Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyama, S., and Akasaka, K. (2006) Cold denaturation of ubiquitin at high pressure, *Magn Reson Chem 44*, S108-S113.

[34] Adrover, M., Esposito, V., Martorell, G., Pastore, A., and Temussi, P. A. (2010) Understanding cold denaturation: The case study of Yfh1, *J Am Chem Soc 132*, 16240-16246.

[35] Adrover, M., Martorell, G., Martin, S. R., Urosev, D., Konarev, P. V., Svergun, D. I., Daura, X., Temussi, P., and Pastore, A. (2012) The role of hydration in protein stability: Comparison of the cold and heat unfolded States of Yfh1, *J Mol Biol 417*, 413-424.

[36] 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.

[37] 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*, 5374-5374.

[38] Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318*, 571-582.

[39] 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.

[40] 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.

[41] Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure, *J Am Chem Soc 132*, 4669-4677.

[42] Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2007) 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.

[43] Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on Unix pipes, *J Biomol NMR 6*, 277-293.

[44] Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol 278*, 313-352.

[45] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36*, 1277-1282.

[46] Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays*, Wiley, New York.

[47] Chan, H. S., Shimizu, S., and Kaya, H. (2004) Cooperativity principles in protein folding, *Methods Enzymol 380*, 350-379.

[48] Barrick, D. (2009) What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding?, *Phys Biol 6*, 015001.

[49] Makhatadze, G. I., Clore, G. M., and Gronenborn, A. M. (1995) Solvent isotope effect and protein stability, *Nat Struct Biol 2*, 852-855.

[50] Kresheck, G. C., Schneider, H., and Scheraga, H. A. (1965) The effect of D_2O on the thermal stability of proteins. Thermodynamic parameters for the transfer of model compounds from H_2O to D_2O, *J Phys Chem 69*, 3132-3144.

[51] Baiz, C. R., Peng, C. S., Reppert, M. E., Jones, K. C., and Tokmakoff, A. (2012) Coherent two-dimensional infrared spectroscopy: Quantitative analysis of protein secondary structure in solution, *Analyst 137*, 1793-1799.

[52] Volkov, V., and Hamm, P. (2004) A two-dimensional infrared study of localization, structure, and dynamics of a dipeptide in membrane environment, *Biophys J 87*, 4213-4225.

[53] Choy, W. Y., Mulder, F. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and Kay, L. E. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques, *J Mol Biol 316*, 101-112.

[54] 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.

Chapter 5

The Denatured State Ensemble Populated at High and Low Temperatures Exhibits Distinct Structural Properties

Abstract

There is debate regarding the degree of compaction of protein denatured state ensembles (DSEs) at high temperatures and their tendency to form secondary structure. Single-molecule Förster resonance energy transfer (smFRET) and small angle X-ray scattering (SAXS) have been used to study the heat-induced DSEs of several proteins, but the results have not led to a consensus. Some smFRET studies have suggested that the DSE becomes more compact at high temperatures, while SAXS studies on other proteins show no decrease in radius of gyration (R_g) as the temperature increases. In contrast to the case with thermally unfolded proteins, relatively few studies have been reported on the temperature dependence of the degree of expansion or contraction of the cold unfolded DSE. In order to simultaneously investigate the properties of the DSEs in the high and low temperature regimes, SAXS measurements were performed on the I98A point mutant of the C-terminal domain of the ribosomal protein L9 (CTL9) under native conditions, pH 5.6, 10 mM DMG (3,3-dimethylglutaric acid) and 120 mM NaCl without denaturant. This mutant significantly populates the folded state and DSE over a wide temperature range, allowing direct comparison of the properties of the thermal and cold induced DSE. Scattering profiles were collected over the temperature range of 5 to 60 °C. I98A-CTL9 exhibits maximum stability at 30°C, and the experimental temperature range covers both cold and heat denaturation, allowing a direct comparison of the two different DSEs for the same protein. The R_g of the unfolded I98A-CLT9 at each temperature point was

estimated by subtracting the contribution from the folded ensemble. The R_g of the DSEs showed no detectable change above 30 °C, but did increase below 30 °C. R_g of the DSE was found to be 27.8 ± 1.7 Å at 5°C, 21.8 ± 1.9 Å at 30°C, and 21.7 ± 2.0 Å at 60°C. These observations differ from some recent smFRET experiments that claim the DSE becomes more compact at high temperature. Temperature dependent circular dichroism (CD) reveals that there is more residual helical structure in the cold unfolded DSE than in the thermal induced DSE, even though the cold unfolded DSE is more expanded. This observation decouples the extend of secondary structure from chain compaction.

Note: The studies presented in this chapter are in preparation for submission. This chapter contains direct excerpts from the draft paper with suggestions and revisions by Prof. Daniel P. Raleigh. I thank Dr. Marc Allaire for the help of SAXS experiments.

5.1. Introduction

Proteins undergo unfolding at both high and low temperatures, this well-known behavior is predicted by the basic thermodynamics of protein stability. The Gibbs-Helmholtz equation defines two transition temperatures, the midpoint of cold denaturation, T_c, and the midpoint of thermal denaturation, T_m. Each is defined as the temperature at which $\Delta G^o(T)$ of unfolding equals zero. As described in the previous chapter, cold denaturation can be rationalized by the decrease in the strength of the hydrophobic effect as temperature is decreased below the temperature of maximum stability (T_s), together with solvation effects including water penetration.[1-10] The mechanism of protein heat denaturation is related to changes from the hydrophobic effect and temperature-dependent solvation free energy and changes in chain entropy. It is of interest to compare the cold and heat induced transition to deduce if they are thermodynamically equivalent, and have similar structural properties. Work with a monomeric construct of λ repressor suggested that the cold and thermally induced DSEs populated in 3 M urea are thermodynamically equivalent and it was observed that the two unfolded ensembles have very similar one-dimensional ^{1}H NMR spectra.[11] Recent studies have shown that the cold unfolded state can expand as the temperature is decreased from T_s, but there are conflicting reports on the behavior of unfolded proteins as the temperature is increased above T_s. Single-molecule Förster resonance energy transfer (smFRET) studies have been interpreted to demonstrate chain compaction as the temperature is increased.[12, 13] In contrast, small angle X-ray scattering (SAXS) measurements argue that the denatured state ensemble (DSE) does not become more compact as the temperature is increased.[14]

Since T_c is often below the freezing point of water, studies of protein cold denaturation are usually conducted by modifying the solution by adding denaturant, increasing pressure, using extreme pH values, or encapsulation in reverse micelles.[15-19] These have largely prevented a direct comparison between the compactness of the heat and cold induced DSEs, under near native like conditions. Ideally one would populate the thermal and cold induced DSE of the protein of interest under the same solution conditions (pH, ionic strength, and pressure) with the least amount of additional perturbation. A few proteins have a T_c above 0 °C, allowing studies of the protein cold denaturation in water without additional perturbations.[14, 20-22]

The effects of temperature on the thermally induced DSE are controversial. Recent smFRET experiments have been interpreted to indicate a continuous compaction of the small cold shock protein from Thermotoga maritime (CspTm) [23], the intrinsically disordered protein prothymosin α (ProTα) [23], and yeast frataxin (Yfh1) as the temperature increases[13]. SAXS studies of other proteins have suggested that the thermally unfolded state does not become more compact as the temperature is increased.[14] The cold denatured ensembles of I98A-CTL9 has been reported to be expanded (see Chapter 4).[24] However, the question of whether the cold unfolded ensemble has a different compactness than the thermal unfolded DSE is still under debate. Differences in the tendency to form secondary structure and the degree of protein hydration in the cold and heat induced DSE of the protein Yfh1 have been reported based on NMR analysis,

however, it is not known if there is a direct relationship between compaction and the level of residual secondary structure.[14]

Temperature dependent SAXS measurements were performed on I98A-CTL9 at pD 5.6 (uncorrected pH reading), in order to investigate the properties of the DSE in the high-temperature regime and to compare it to the cold unfolded DSE. D_2O was used to facilitate comparison with earlier studies. Scattering profiles were collected over the temperature range of 5 to 60 °C. I98A-CTL9 exhibits maximum stability at 30°C, thus this temperature range covers both cold and heat induced denaturation, and allows us to direct compare the two different DSEs using the same protein. SAXS studies were complimented by temperature dependent CD studies to probe residual secondary structure. The results were compared to published studies on another protein, yeast frataxin (Yfh1)[14].

5.2. Materials and methods

5.2.1. Mutagenesis, protein expression, and purification

I98A-CTL9 was expressed and purified using as previously described.[24-28] No protease inhibitors were added during the purification. The protein was purified using a SP-Sepharose fast flow ion-exchange column (GE Healthcare). 20 mM Tris buffer at pH 7.4 was used as the equilibrium buffer, and the protein was eluted with a 0 to 2 M NaCl gradient. Proteins were further purified with the reverse phase HPLC using a C8 column (Vydac). An A-B gradient was used, buffer A contained 99.9% H_2O and 0.1% TFA

(trifluoroacetic acid) and buffer B contained 90% acetonitrile, 9.9% H_2O, and 0.1% TFA. The protocol used is 100% A for 10 mins after injecting the sample followed by 25-55% B in 60 mins. I98A-CTL9 eluted around 40-42% B. The identity of the protein was confirmed by DNA sequencing and MALDI-TOF mass spectroscopy. The observed molecular weight was 9940.9 ± 1.2 Da and the expected molecular weight was 9939.5 Da. The yield of the protein was 70-80 mg/L in LB media. A similar yield was observed for wild type CTL9. The purity was tested by analytical HPLC.

5.2.2. Circular dichroism (CD) spectroscopy

CD experiments were performed using a Chirascan CD spectrometer. The protein was dissolved in 10 mM DMG (3,3-dimethylglutaric acid) and 120 mM NaCl buffer in D_2O, at a protein concentration of about 20 μM, at pD 5.6 (uncorrected pH reading). D_2O was used to allow comparison with ^1H NMR studies. DMG was chosen because of its small heat of ionization which helps maintain a constant pD (pH) value over a wide temperature range.

Far-UV CD spectra of I98A-CTL9 and wild type CTL9 in buffer were collected in a 1 mm cuvette from 196 nm to 260 nm with a 0.2 nm increment and averaged with 3 repetitions. Wild type CTL9 is folded over the temperature range studied here and provides a baseline for the folded state CD signal. A sample of I98A-CTL9 in 8 M urea was used to provide a completely unfolded baseline. This spectrum was collected in the same cuvette over the wavelength range from 210 nm to 260 nm. The population-

weighed CD spectrum of the folded state was subtracted from the experimental I98A-CTL9 CD spectrum at pD 5.6 in order to obtain the unfolded I98A-CTL9 CD spectrum, using the equation below:

$$\theta_{I98A,\,unfolded} = (\theta_{I98A} - p_{folded} \cdot \theta_{CTL9wt}) / (1 - p_{folded}) \qquad \text{(eq 5-1)}$$

where θ_{I98A} is the CD signal collected for I98A-CTL9 at pD 5.6, containing both the folded and unfolded ensembles; θ_{CTL9wt} is the CD signal collected for the wild type CTL9 at pD 5.6, only the folded state is present in this sample, providing a spectrum of the folded state for the subtraction; p_{folded} is the fractional population of folded I98A-CTL9 present under the conditions of the CD measurements. p_{folded} was determined by thermal denaturation monitored by CD; $\theta_{I98A,unfolded}$ is the calculated I98A-CTL9 DSE CD signal.

The mean residue ellipticity at 222 nm, calculated for the I98A-CTL9 DSE ($[\theta]_{DSE}$), was used for the determination of the fraction helix, following

$$f_h = ([\theta]_{DSE} - [\theta]_C)/([\theta]_H - [\theta]_C) \qquad \text{(eq 5-2)}$$

where $[\theta]_C$ corresponds to the signal for a random coil and $[\theta]_H$ corresponds to the signal expected for a 100% helix, calculated from

$$[\theta]_H = -40,000 \times (1 - 2.5/n) + 100 \times T \qquad \text{(eq 5-3)}$$

$$[\theta]_C = 640 - 45 \times T \qquad \text{(eq 5-4)}$$

Where n is the number of the residues in the peptide, and T is the temperature (°C).[29]

Thermal denaturation experiments were conducted as a function of pD between 4.0 and 8.0. Each thermal denaturation was performed by monitoring the ellipticity at 222 nm using a 1 cm cuvette, from 4 ˚C to 98 ˚C, with a 2 ˚C step and a heating rate of 1 ˚C/min.

The pD 4.0 thermal denaturation data was fit to a quadratic equation to obtain the unfolded state signal as a function of temperature, I98A-CTL9 has been shown to populate the DSE for all temperatures at this pD.[24, 27] The completely folded baseline was estimated from fitting thermal denaturation data for I98A-CTL9 at pD 8.0, and extrapolating the folded state signal. Data was fit to the following equation to obtain thermodynamic parameters for unfolding. The observed CD signal, θ(*T*), was fit to the equation:

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

where $\Delta G_u^\circ(T)$ is the free energy change upon thermal unfolding described by the Gibbs-Helmholtz equation:

$$\Delta G_u^\circ(T) = \Delta H^\circ(T_m) - T\Delta H^\circ(T_m)/T_m + \Delta C^\circ_P [T - T_m - T \ln(T/T_m)] \qquad \text{(eq 5-6)}$$

a_n, b_n, a_d and b_d are parameters which define the signals of the native state (N) and denatured state ensembles (D) as a function of temperature. T_m is the thermally induced unfolding midpoint temperature. $\Delta H^\circ(T_m)$ is the enthalpy change at T_m. ΔC°_P is the change in heat capacity upon unfolding. The signal expected for a fraction folded of 0.5 was estimated by taking the average of the native state baseline and the unfolded state baseline as a function of temperature.

Urea unfolding data were collected at 25°C on wild type CTL9 and the I98A-CTL9 mutant, using an AVIV Instruments model 202SF CD instrument. The native protein sample was titrated with urea in 0.25 M increments until the final urea concentration reached 10 M. The buffer used was the same one used for the thermal unfolding

experiments, and the urea concentration was determined by measuring the refractive index.

5.2.3. Small angle X-ray scattering measurements

Samples of I98A CTL9 were prepared in buffer consisting of 10 mM DMG and 120 mM NaCl in 100 % D_2O, with the pD adjusted to 5.6. Scattering experiments were performed at beamline X9 at Brookhaven National Laboratory, National Synchrotron Light Source I (Upton, New York, USA). Protein samples were injected into a 1 mm capillary continuously at a rate of 0.67 μL/s in order to avoid radiation damage. The exposure time for each measurement was 30 s. Scattering data were collected for I98A CTL9 at a protein concentration of 3.75 mg/mL, from 5 to 60 °C, with an increment of 5 °C. Each sample was measured three times and then averaged before data analysis. A sample of wild type CTL9 was analyzed using the same sample preparation method to provide the folded scattering profile at each temperature point. The scattering profiles of the I98A-CTL9 DSE were obtained after subtracting the population-weighted folded signal, using:

$$I_{\text{I98A, unfolded}} = (I_{\text{I98A}} - p_{\text{folded}} \cdot I_{\text{CTL9wt}}) / (1 - p_{\text{folded}}) \qquad \text{(eq 5-7)}$$

where I_{I98A} is the SAXS data collected for the sample of I98A-CTL9 at pD 5.6, containing both the folded and unfolded ensembles; I_{CTL9wt} is the scattering profile collected for wild type CTL9 at pD 5.6, only the folded state is populated under this condition providing a folded scattering profile for the subtraction; p_{folded} is the population of folded I98A-CTL9 present under the conditions of the SAXS measurements; $I_{\text{I98A, unfolded}}$ is the calculated I98A-CTL9 DSE scattering profile.

The program pyXS (http://www.bnl.gov/ps/x9/software/pyXS.asp) was used for buffer subtraction. The radius of gyration (R_g) was obtained based on the scattering profiles using the Guinier approximation using the program PRIMUS,[30]

$$I(q) = I(0) \bullet \exp(-R_g^2 q^{2/3}) \qquad \text{(eq 5-8)}$$

where $I(q)$ is the intensity at scattering vector q.[31]

A Guinier analysis was performed directly on I98A-CTL9 scattering data in order to obtain the apparent R_g values, $R_{g,\,app}$, at each temperature point. The value of the R_g for the DSE, $R_{g,\,DSE}$ can then be estimated from:

$$R_g^2{}_{app} = p_{folded} \cdot R_g^2{}_{folded} + p_{DSE} \cdot R_g^2{}_{DSE} \qquad \text{(eq 5-9)}$$

where p_{folded} and p_{DSE} are the fractional populations of the folded ensemble and the DSE, respectively.

5.3. Results

5.3.1. Thermal denaturation monitored by CD reveals conditions where both heat and cold denaturation of I98A-CTL9 can be studied

The CD signal monitored at 222 nm (**Figure 5-1**) showed that I98A-CTL9 exhibits pH (pD) dependent unfolding over the temperature range studied in DMG buffer. DMG was chosen because of its small heat of ionization[32], which helps maintain a constant pH (pD) over the temperature range studied, allowing the direct comparison of the heat and cold denatured I98A at the same pD. At pD 5.6 (uncorrected pH-meter reading), equal

amounts of the folded and unfolded ensembles are populated at 30 °C. 30 °C is also the temperature of maximum stability, T_s, for I98A-CTL9 under these conditions. The plot reveals both cold and heat induced unfolding. At 5 °C the population of the folded state is decreased to 13%. A similar population of the folded state is reached at a temperature close to 60 °C during the thermal unfolding. The calculated folded and unfolded baselines are shown in **Figure 5-1A**, allowing an estimation of the fractional folded population at any temperature and pD. The estimated fraction folded at each temperature point where SAXS data were collected is listed in **Table 5-1** and shown in **Figure 5-1B**.

5.3.2. The R_g of the I98A-CTL9 DSE shows different trends at low and high temperature

We first examined the concentration dependence of the scattering profile. I98A-CTL9 showed no significant changes in the scattering profile at 7 °C at three different concentrations (2.5, 3.75, and 5 mg/mL) (**Figure 5-2A**). Guinier analysis (**Figure 5-2B**) showed similar apparent R_g values for these three protein concentrations. The R_g values of the DSE are listed in **Table 5-1**. There were no concentration-dependent deviations observed. Scattering data for wild type CTL9 were collected at pD 5.6 over the temperature range of 5 to 60 °C. In this temperature range, wild type CTL9 is 100% folded, and Guinier analysis (**Figure 5-3A**) revealed a consistent R_g values of 15 Å for the protein (**Table 5-2**). Scattering profiles for the I98A-CTL9 at pD 5.6 were collected over the temperature range of 5 to 60 °C.

The apparent (observed) R_g values obtained by applying the Guinier approximation directly to the scattering data (**Figure 5-3B**), showed an increase in the apparent R_g at both low and high temperature, but the increase at the high temperature end is not as pronounced as that in the low temperature regime (**Figure 5-4**). At T_s (30 °C), the apparent R_g is 19 Å, a 26 % increase compared with wild type CTL9 (folded state control, R_g of 15 Å) at the same temperature. The observed apparent R_g at 5 °C, showed an increase of 32% to 25 Å, relative to the value at 30 °C. At 55 °C, the observed R_g only increased by 1-2 Å, reaching a value of 21 Å. The apparent R_g values are listed in **Table 5-2**. The data demonstrate that the cold and heat unfolded ensembles behave differently. These differences are more apparent once the contribution of the native state is deconvolved.

In order to estimate the R_g values of the I98A-DSE, $R_{g,\,DSE}$, the signal due to the folded state was subtracted from the observed I98A-CTL9 scattering profile. The folded state scatting profile is calculated from the scattering data of wild type CTL9 at each temperature, together with the estimated fraction folded of I98A-CTL9 at each temperature. The fraction folded at each temperature point is listed in **Table 5-2** and plotted as a function of temperature in **Figure 5-5**. The R_g of the I98A-CTL9 DSE was determined from the resulting scattering profile (**Figure 5-3C**). The $R_{g,\,DSE}$ at T_s is 22 Å. $R_{g,\,DSE}$ showed a clear increase for the cold-induced DSE, with an increase of 85% relative to the folded state R_g value, to 28 Å at 5 °C, while the R_g of the heat-induced DSE stayed constant at ~ 22 Å, a 47 % increase from the folded R_g value (**Figure 5-5**). In the low temperature induced unfolding, between 5 and 30 °C, a continuous increase in $R_{g,}$

$_{DSE}$ was observed, in contrast, at temperatures higher than T_s, the $R_{g,\,DSE}$ was constant. We further analyzed the SAXS data of I98A-CTL9 using a two-state system estimation, using equation 5-9. The calculated $R_{g,\,DSE}$ values are listed in **Table 5-2**, and are in good agreement with the $R_{g,\,DSE}$ determined by the previous analysis.

The pKa of buffer may change at high temperatures and in order to test this effect on the analysis of the I98A-CTL9 DSE scattering data, the pH (pD) dependence of I98A-CTL9 at 60 °C was also examined at three different pD values, 5.0, 5.6, and 6.2. The analysis of scattering data (**Figure 5-6**) is shown in **Table 5-3**. There are no significant effects of a shift to lower pD value from pD 5.6, while a slightly more obvious changes were observed for the pD value increses by 0.6 unit.

5.3.3. The I98A-CTL9 DSE contains different amounts of α-helical structure at low and high temperatures

In order to compare the α-helical content of the cold- and heat- induced DSE of I98A-CTL9, far UV CD spectra were recorded as a function of temperature (**Figure 5-7**, **5-8**, **5-9**, and **5-10**). Wild type CTL9 in buffer was used to provide the spectrum of the folded state. Wild type CTL9 undergoes thermally unfolding with a T_m of 78.4 °C, thus it can used to provide a folded baseline for the temperature range of this study. The folded spectrum, together with the estimated population of the folded ensemble for I98A-CTL9, can be used to calculate the CD signal from the DSE of I98A-CTL9 using equation 5-1. The fractional folded population is calculated to be ~ 27% at 15 and 50 °C and ~ 12% at

5 and 60 °C (**Table 5-2**). The two temperatures allow us to compare the heat and cold induced DSE under conditions where a similar folded population is present, i.e., at points where ΔG^{o}_{u} is equal. Cold-denatured I98A-CTL9 keeps α-helical content 11% and 9% at 5 and 15 °C (**Figure 5-7** and **Figure 5-8**) than the heat-denatured DSEs, for which the estimated helical content is 3% and 7% at 60 and 50 °C (**Figure 5-9** and **Figure 5-10**).

A sample of I98A-CTL9 in 8 M urea provides another set of random coil reference CD spectra. A urea titration experiments on I98A-CTL9, as monitored by the CD signal at 222 nm, is shown in **Figure 5-11**. At 25 °C, I98A-CTL9 is unfolded above urea concentrations of 4 M. Thus 8 M urea provides a reference for the unfolded state. CD spectra were recorded in 8 M urea at 5, 15, 50, and 60 °C. The percentage helical content was estimated to be 9 and 6% for cold denatured I98A-CTL9 at 5 and 15 °C using the unfolded 8 M urea reference (**Figure 5-7** and **Figure 5-8**). The heat induced DSE of I98A-CTL9 showed a very similar CD signal to that of the urea unfolded state (**Figure 5-9** and **Figure 5-10**), with an estimation of helical content of only 0.5% and 4% at 60 and 50 °C, respectively. Both of the calculations show that there is less α-helical structure in the heat-induced DSE of I98A-CTL9, relative to the cold unfolded DSE.

5.3.4. I98A-CTL9 in denaturant has a R_g value consistent with a random coil model

The expanded R_g values of the I98A-CTL9 populated in 6 M GdnHCl (**Figure 5-12**) from the Guinier analysis showed good agreement with the R_g value predicted for a random coil in good solvent, which is ~29 Å [33] , the experimental value is 30 ± 1.5 Å. All the R_g values of the I98A-CTL9 DSE at low and high temperatures are smaller than this. In

addition, the R_g values of the I98A-CTL9 populated in 8 M urea are 31.7 ± 1.4 Å, 31.3 ± 1.8 Å, and 31.5 ± 0.9 Å at 5, 30, and 50 °C, respectively. This set of data further confirmed the R_g values of the I98A-CTL9 DSE in denaturant are consistent with a random coil model.

5.4. Discussion

The results of the CD studies show that the α-helical content is different for the cold- and heat-induced DSE, with less detectable secondary structure in the high-temperature regime. This is different from the behavior reported for some IDPs, which have been reported to undergo high-temperature induce refolding based on, for example, human α-synuclein and the phosphodiesterase γ-subunit over the temperature range of 3 to 30 or 3 to 50 °C, respectively.[34] The activation domain of the activator for the cytosolic C-terminal distal tail of the human sodium-proton exchanger 1 (hNHE1cdt), thyroid hormone and retinoid receptors (ACTR), and the S-shape delayed protein (Spd1) also have been reported to display an increase in secondary structure content in the high temperatures range up to 95 °C.[34] However, one should bear in mind that caution needs to be paid when analyzing CD spectra at higher temperatures, because the interpretation of changes in the spectra can be ambiguous. Kjaergaard et al suggested that the structural changes in one segment of a protein may be cancelled out by the signal contributed by another segment within the ensemble. In this case, the seemingly increased helical structural element in the CD spectra at high temperatures is not actually induced by the formation of transient helices. Secondary chemical shift analysis by NMR has shown that

there is even a loss in transiently formed α-helices at high temperature.[34] The net conclusion is that the CD changes are caused primarily by the redistribution of the statistical coil ensembles, thus the interpretation of results from bulk methods should be carried out with caution. The work reported here reveals a decrease in the α-helical content as the temperature is increased.

Some heat-induced DSEs and IDPs at higher temperatures have been reported to be more compact at higher temperatures, relative to the DSE populated near room temperature, in contrast to the behavior expected for a polymer chain with temperature independent monomer-monomer interactions.[35, 36] Uversky suggested that the increased strength of hydrophobic interactions, which leads to a stronger hydrophobic attraction at elevated temperatures, might be the reason for the refolding of IDPs.[37]

Our data shows that the DSE does not become more compact as the temperature is increased above T_s. Cold and heat induced denaturation has been examined for Yfh1 by both smFRET and SAXS.[13, 14] Based on SAXS studies, it has been concluded that the cold and heat induced DSEs of Yfh1 have subtle, but clear differences in compactness and size-distribution, in spite of having structurally equivalent local structure. Both the cold and heat induced Yfh1 showed an obvious increase in apparent R_g relevant to the value at 20 °C, however that work did not report a value for the DSE.[14] The results we observed for the apparent R_g of I98A-CTL9 agree with the reported behavior for the Yfh1 at high temperatures, but our deconvolved data shows that the R_g of the DSE is independent of the temperature. At 0 °C, Yfh1 showed a bigger population of more

compact structures, while at 50 °C, the R_g distribution shifted to resemble a random coil. Molecular dynamics (MD) simulations of Yhf1 further suggested that different degrees of hydration are observed for the cold and heat induced DSEs, with a higher degree of hydration found for the cold denatured state.[13]

Different results for the dimension of the heat denatured DSE of Yfh1 have been reported based on smFRET and SAXS.[13, 14] There are likely several factors that contribute to the discrepancy. The differences may be due to dynamics at elevated temperatures which occurs on the time scale of the single molecule measurement. This can lead to compact structures being over averaged. In addition, smFRET studies require large dyes, relatively hydrophobic dyes, which may introduce perturbations to the system of interest. Last but not least, the R_g values are not directly obtained from the smFRET data, but depend on the use of models to fit the data and the value of R_g can be affected by the choice of the model.

Finally, it is interesting to note that our data reveals that there is no direct relationship between the residual α-helical content and the compactness of the DSE, as judged by R_g. This may seem counter-intuitive, but recent work has shown the helical structure is fully compatible with a "random-coil" R_g.[38-40]

Table 5-1. R_g values of the DSE of I98A-CTL9 as a function of protein concentration at 7 °C, in 10 mM MOPS and 150 mM NaCl, at pH 6.0 (in H_2O).

[I98A-CTL9] (mg/mL)	$R_{g,\ app}$ (Å)[a]	$R_{g,\ DSE}$ (Å)[b]	$R_{g,\ DSE}$ (Å)[c]
2.5	21.8 ± 0.7	27.1 ± 1.4	28.5 ± 1.2
3.75	21.6 ± 0.9	28.4 ± 0.9	28.2 ± 1.4
5	22.1 ± 1.0	27.8 ± 1.6	29.1 ± 1.5

a) Apparent R_g values obtained directly by the Guinier analysis from the observed I98A-CTL9 scattering profile; *b*) R_g values of the I98A-CTL9 DSE obtained by Guinier analysis on the DSE scattering profile after subtracting the signal from the weighted folded state contribution.; *c*) R_g values of the I98A-CTL9 DSE calculated by the two-state equation (5-9). The error bars represent the linear fitting in the Guinier approximation.

Table 5-2. R_g values and fraction folded of I98A-CTL9, in 10 mM DMG and 120 mM NaCl, 100% D_2O.

Temperature (°C)	$R_{g,\ app}$ (Å)[a]	$R_{g,\ folded}$ (Å)[b]	p_{folded} (%)[c]	$R_{g,\ DSE}$ (Å)[d]	$R_{g,\ DSE}$ (Å)[e]
5	26.0 ± 1.7	14.9 ± 0.3	13.2	27.3 ± 1.7	27.8 ± 1.7
7	24.4 ± 1.4	15.1 ± 0.4	14.8	25.7 ± 1.5	27.0 ± 1.6
10	23.7 ± 1.4	15.1 ± 0.4	16.3	25.0 ± 1.5	26.5 ± 1.6
12	24.0 ± 1.9	15.1 ± 0.4	21.2	25.9 ± 1.6	26.6 ± 1.3
15	22.3 ± 1.2	15.2 ± 0.4	26.2	24.3 ± 1.4	24.1 ± 1.5
20	20.5 ± 1.4	15.1 ± 0.4	36.4	23.0 ± 0.9	22.7 ± 1.4
25	19.7 ± 1.6	15.1 ± 0.3	44.9	22.8 ± 1.0	22.8 ± 1.8
30	18.7 ± 1.4	14.9 ± 0.3	50.7	21.9 ± 0.9	21.8 ± 1.9
35	18.6 ± 1.2	14.8 ± 0.4	50.9	21.9 ± 0.8	20.0 ± 2.1
40	19.3 ± 1.6	15.0 ± 0.3	46.8	22.4 ± 1.0	22.6 ± 1.9
45	20.0 ± 1.5	15.0 ± 0.3	38.0	22.5 ± 0.9	20.7 ± 1.8
50	19.6 ± 1.8	14.9 ± 0.4	28.1	21.2 ± 1.1	21.4 ± 1.7
55	21.4 ± 1.8	14.8 ± 0.3	18.0	22.6 ± 1.2	21.9 ± 1.7
60	20.8 ± 1.6	15.2 ± 0.2	12.1	21.5 ± 0.9	21.7 ± 2.0

a) Apparent R_g values obtained directly by the Guinier analysis from the observed I98A-CTL9 scattering profile; *b*) R_g values of wild type CTL9, used as a control for the folded state over the temperature range used for this study; *c*) the population of the folded ensemble for I98A-CTL9 as a function of temperature, estimated from the CD monitored thermal denaturation data; *d*) R_g values of the I98A-CTL9 DSE calculated by the two-

state equation (5-9); *e*) R_g values of the I98A-CTL9 DSE obtained by Guinier analysis on the DSE scattering profile after subtracting the population weighted signal from the folded state contribution. The error bars represent the linear fitting in the Guinier approximation.

Table-5-3. The pD-dependence of the R_g values of I98A-CTL9 DSE at 60 °C, in 10 mM DMG and 120 mM NaCl, 100% D_2O.

pD	Folded %	$R_{g,\,app}$ (Å)[a]	$R_{g,\,DSE}$ (Å)[b]	$R_{g,\,DSE}$ (Å)[c]
5.0	3.0	25.5 ± 1.9	24.0 ± 1.6	25.8 ± 1.2
5.6	12.1	25.0 ± 2.1	23.4 ± 1.9	26.1 ± 1.3
6.2	34.2	20.7 ± 1.8	21.7 ± 2.1	22.8 ± 1.4

a) Apparent R_g values obtained directly by the Guinier analysis from the observed I98A-CTL9 scattering profile; *b)* R_g values of the I98A-CTL9 DSE obtained by Guinier analysis on the DSE scattering profile after subtracting the signal from the weighted folded state contribution.; *c)* R_g values of the I98A-CTL9 DSE calculated by the equation describing a two-component system.

Table 5-4. Helical fraction estimated for the I98A-CTL9 DSE using the mean residue ellipticity monitored at 222 nm by CD, in 10 mM DMG and 120 mM NaCl, 100% D_2O.

Temperature (°C)	p_{folded} [a]	Helix (%) [b]	Helix (%) [c]
5	13.2	11	9
15	26.2	9	6
50	28.1	7	4
60	12.1	3	0.5

a) Fraction of the folded ensemble at a given temperature; *b*) Estimated by equation 5-2, using $[\theta]_C$ calculated from equation 5-4; *c*) Estimated by equation 5-2, using the $[\theta]_{222\ nm}$ from the I98A-CTL9 in 8 M urea as $[\theta]_C$.

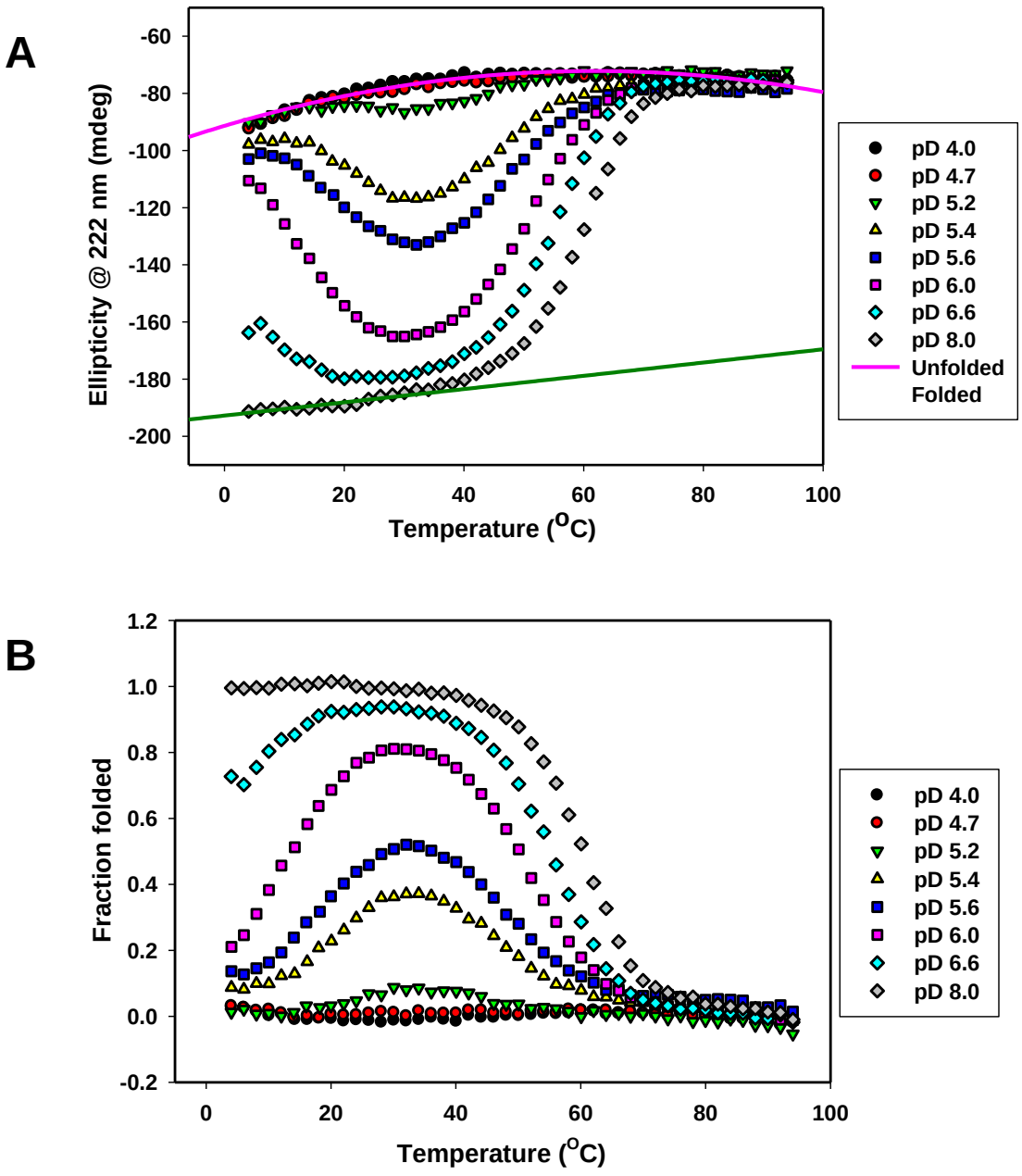

A
Ellipticity @ 222 nm (mdeg)
-60
-80
-100
-120
-140
-160
-180
-200
0
20
40
60
80
100
Temperature (°C)
pD 4.0
pD 4.7
pD 5.2
pD 5.4
pD 5.6
pD 6.0
pD 6.6
pD 8.0
Unfolded
Folded
B
Fraction folded
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
20
40
60
80
100
Temperature (°C)
pD 4.0
pD 4.7
pD 5.2
pD 5.4
pD 5.6
pD 6.0
pD 6.6
pD 8.0

Figure 5-1. (A) CD spectra of I98A-CTL9 at different pD values (uncorrected pH-meter readings), in 10 mM DMG and 120 mM NaCl, 100% D_2O. The estimated folded (green) and unfolded (pink) baselines are shown. The folded baseline was determined by an extrapolation of the pre-unfolding transition of the thermal denaturation curve at pD 8.0. The unfolded baseline was determined by a quadratic fit to the pD 4.0 data. The fraction folded at a given temperature and pD value can be estimated from these plots. (B) Plots of fraction folded as a function of temperature at different pD values.

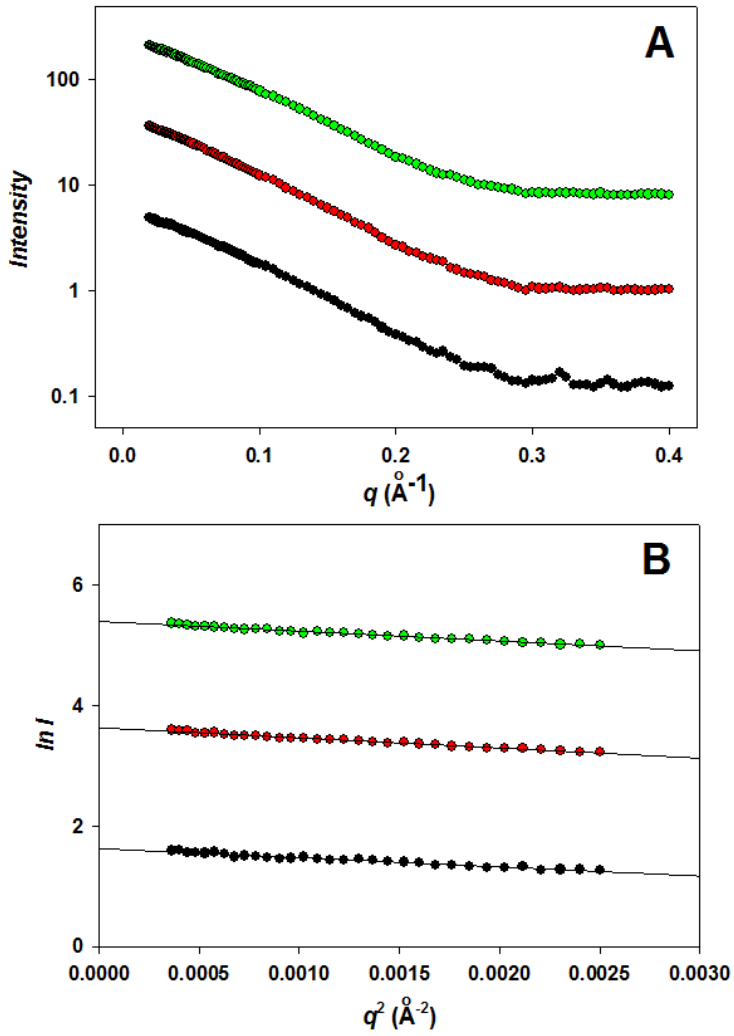
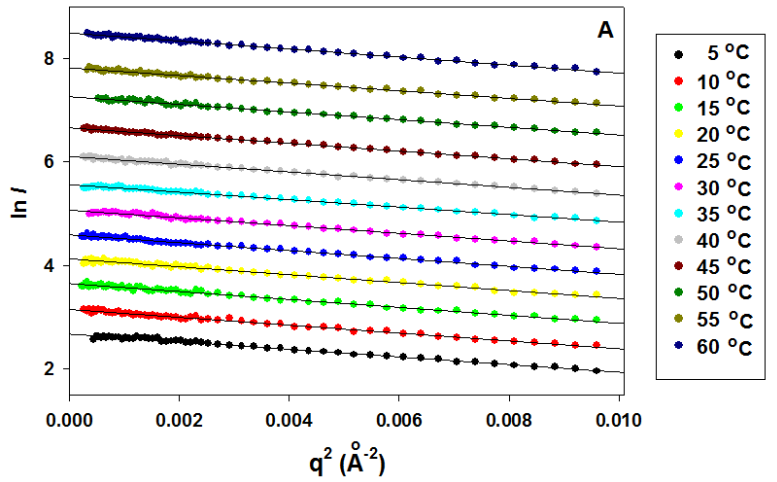
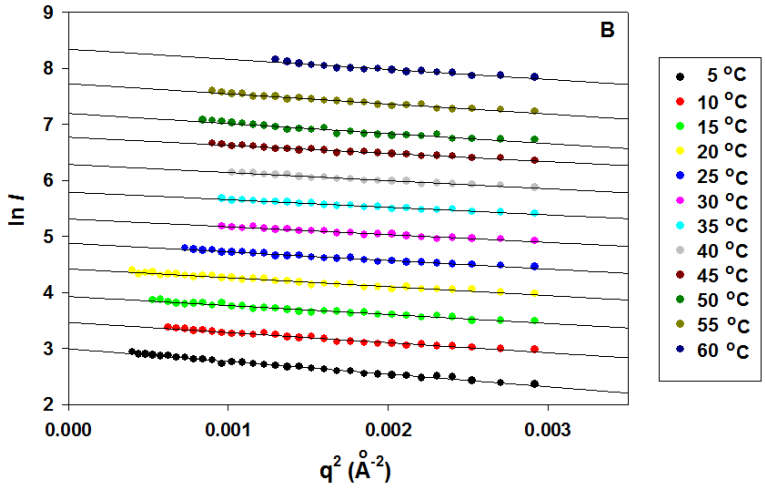


Figure 5-2. (A) Concentration dependence of the scatting profiles of I98A-CTL9 at 7 °C. The protein concentration was 2.5 mg/mL (black), 3.75 mg/mL (red), and 5 mg/mL (green). The buffer contains 10 mM MOPS and 150 mM NaCl at pH 6.0 (in H_2O). The curves are offset for clarity. The identical shape of the curves consistent with the samples being monomeric. (B) Guinier analysis of the data displayed in panel (A). The plots are offset for clarity. The calculated apparent R_g values are 21.9 ± 0.8, 22.5 ± 0.5, 22.1 ± 0.9 Å at 2.5, 3.75, and 5 mg/mL, respectively.


A
ln I
q² (Å⁻²)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C


B
ln I
q² (Å⁻²)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C

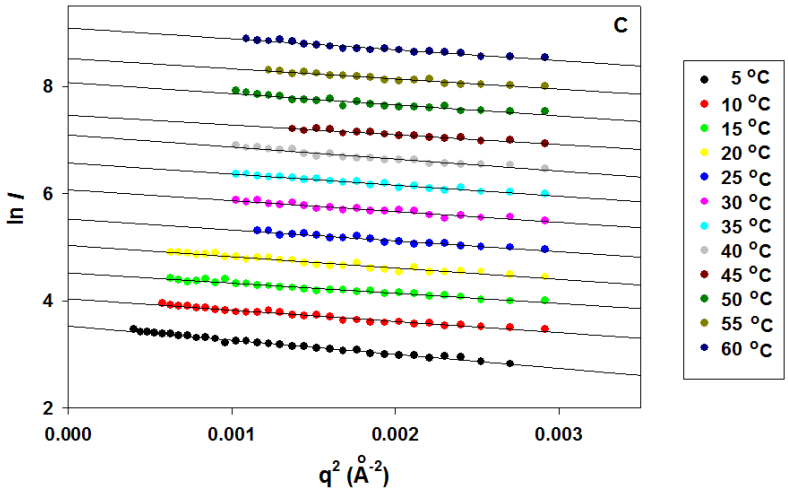

C
ln I
q² (Å⁻²)
5 °C
10 °C
15 °C
20 °C
25 °C
30 °C
35 °C
40 °C
45 °C
50 °C
55 °C
60 °C

Figure 5-3. Temperature dependent analysis of I98A-CTL9 DSE by SAXS. Guinier plots of (A) wild type CTL9 control scattering data, (B) the I98A-CTL9 observed scattering profile before subtracting the folded contribution, and (C) the I98A-CTL9 DSE SAXS scattering profile from 5 to 60 °C, in 10 mM DMG and 120 mM NaCl, 100% D_2O at pD 5.6. The curves are off-set relative to each other for clarity.

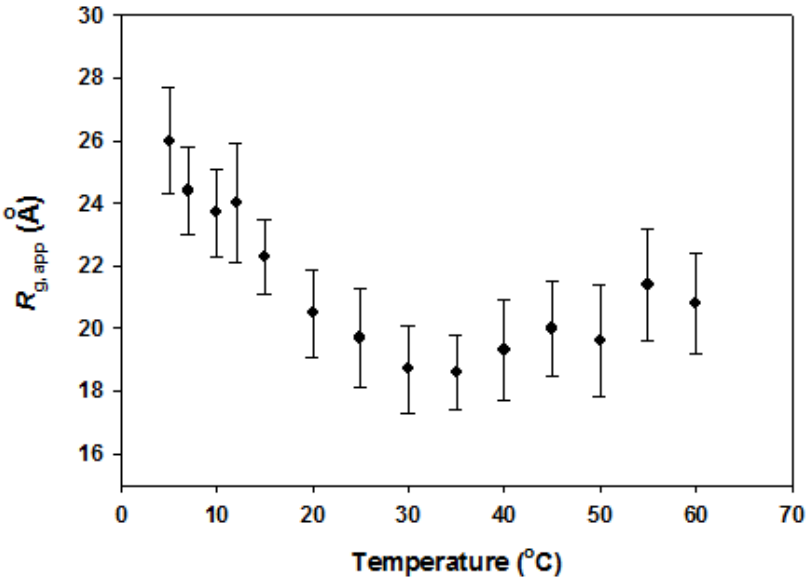


Figure 5-4. Apparent R_g ($R_{g,\ app}$) values at different temperatures from SAXS studies of I98A-CTL9, in 10 mM DMG and 120 mM NaCl, 100% D_2O, obtained by the Guinier analysis on the I98A-CTL9 scattering data at each temperature point. $R_{g,\ app}$ values were calculated from a Guinier analysis of the observed I98A-CTL9 SAXS data. The error bars represent the linear fitting in the Guinier approximation.

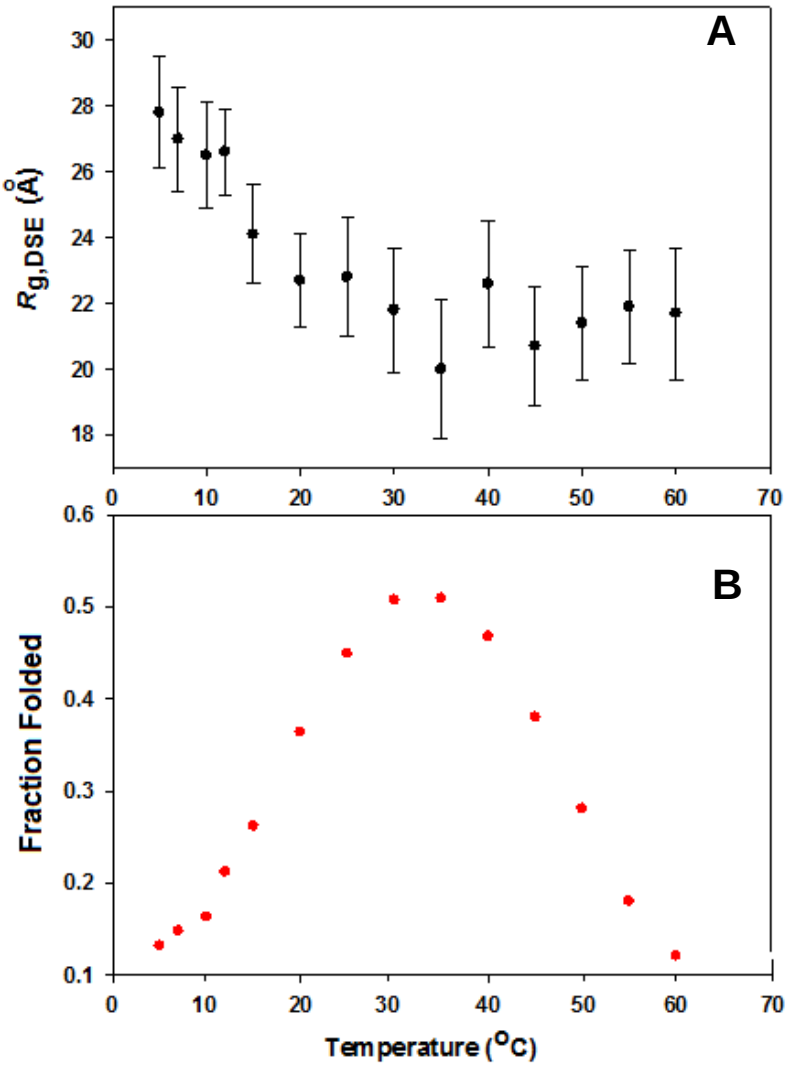


Figure 5-5. (A) $R_{g,\,DSE}$ of I98A-CTL9 as a function of temperature, in 10 mM DMG and 120 mM NaCl, 100% D_2O. Error bars were determined from the linear fitting in the Guinier analysis. (B) A plot of the fraction of folded for I98A-CTL9 over the same range of temperature as SAXS studies.

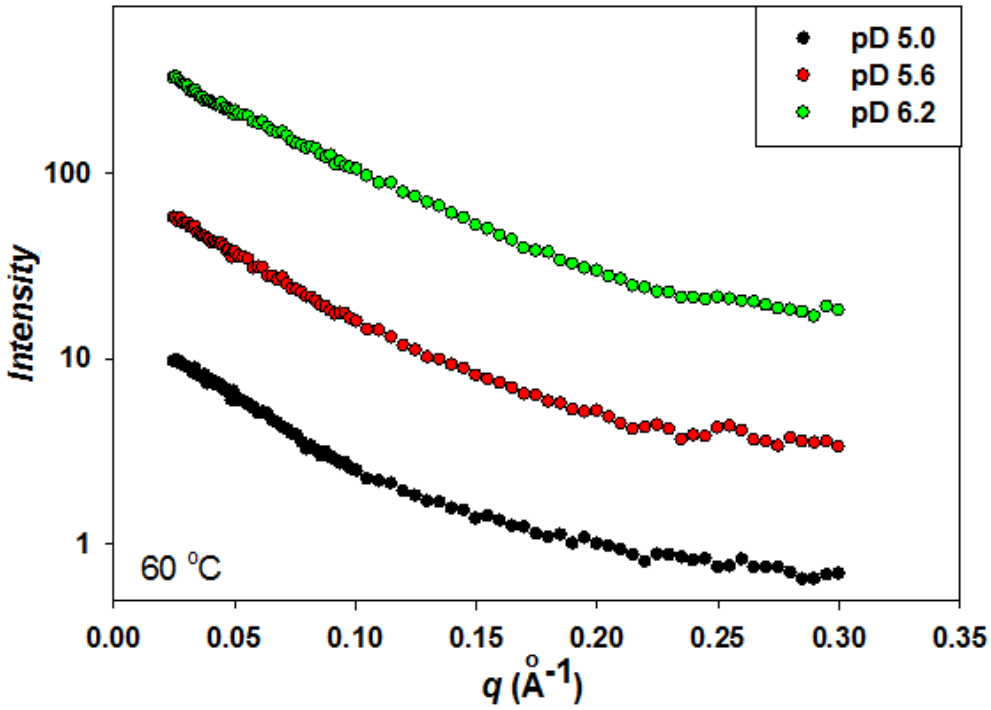


Figure 5-6. Scattering profiles of I98A-CTL9 at 60 °C. Data were collected at three pD values: 5.0, 5.6, and 6.2. The buffer consists 10 mM DMG and 120 mM NaCl, in 100% D_2O.

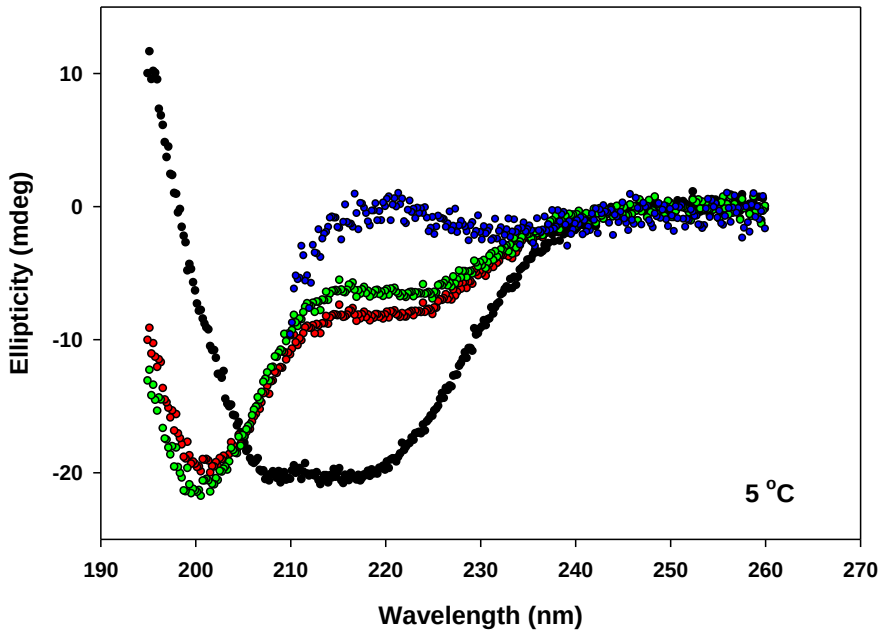


Figure 5-7. CD spectra at 5 °C of I98A-CTL9 (red) and wild type CTL9 (black) in 10 mM DMG and 120 mM NaCl, 100% D_2O at pD 5.6 (uncorrected pH-meter reading). The CD spectrum of the DSE of I98A-CTL9 (green) was calculated by subtracting the population-weighted CD signal of the folded state (using the CD spectrum of wild type as the folded CD spectrum). The CD spectrum of I98A-CTL9 in 8 M urea (blue) is shown for comparison.

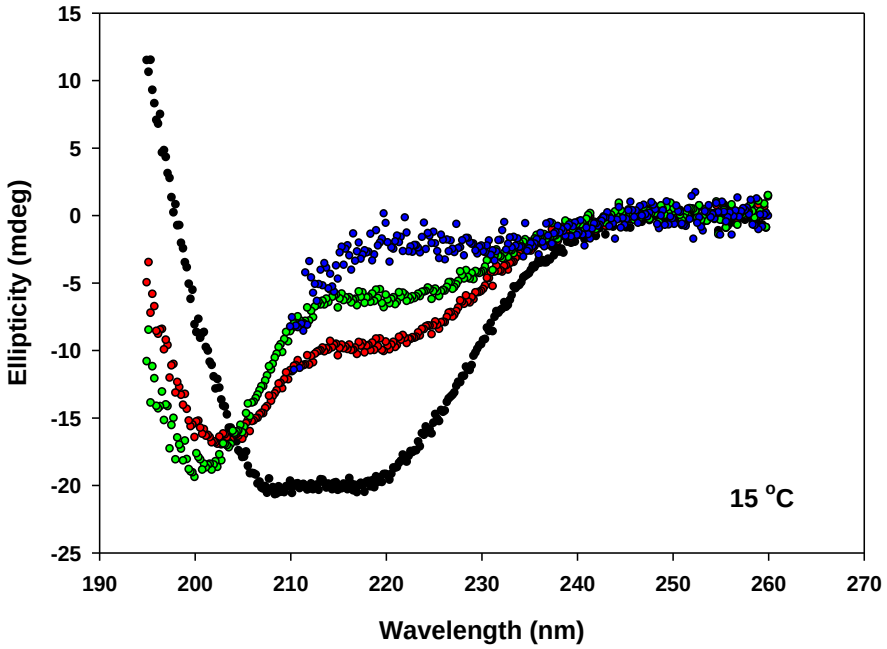


Figure 5-8. CD spectra at 15 °C of I98A-CTL9 (red) and wild type CTL9 (black) in 10 mM DMG and 120 mM NaCl, 100% D_2O at pD 5.6 (uncorrected pH-meter reading). The CD spectrum of the DSE of I98A-CTL9 (green) was calculated by subtracting the population-weighted CD signal of the folded state (using the CD spectrum of wild type as the folded CD spectrum). The CD spectrum of I98A-CTL9 in 8 M urea (blue) is shown for comparison.

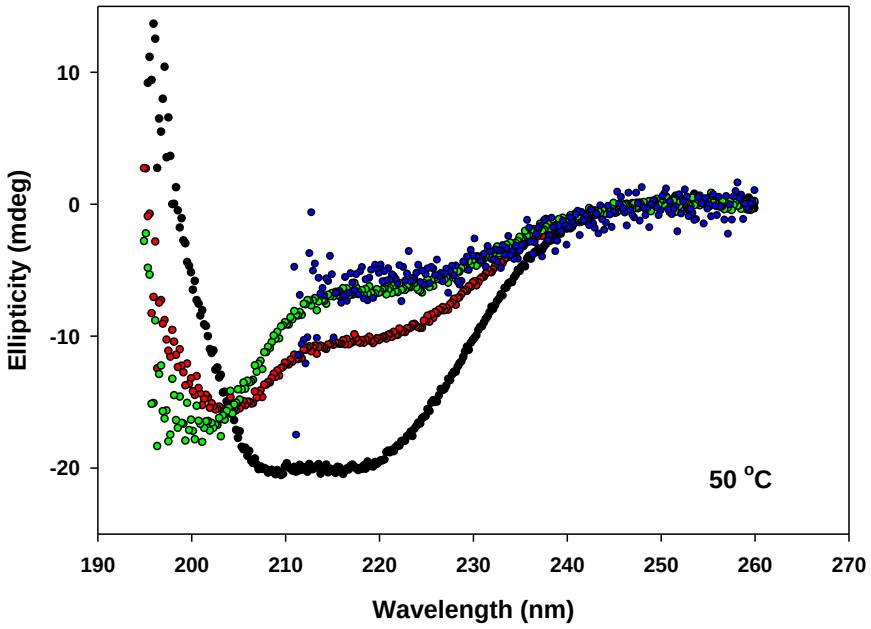


Figure 5-9. CD spectra at 50 °C of I98A-CTL9 (red) and wild type CTL9 (black) in 10 mM DMG and 120 mM NaCl, 100% D_2O at pD 5.6 (uncorrected pH-meter reading). The CD spectrum of the DSE of I98A-CTL9 (green) was calculated by subtracting the population-weighted CD signal of the folded state (using the CD spectrum of wild type as the folded CD spectrum). The CD spectrum of I98A-CTL9 in 8 M urea (blue) is shown for comparison.

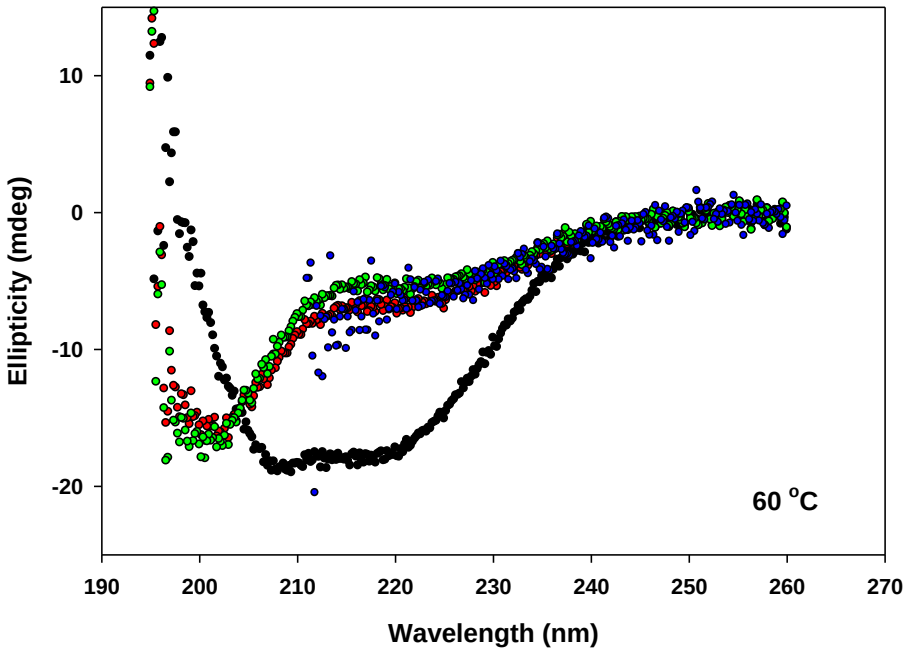


Figure 5-10. CD spectra at 60 °C of I98A-CTL9 (red) and wild type CTL9 (black) in 10 mM DMG and 120 mM NaCl, 100% D_2O at pD 5.6 (uncorrected pH-meter reading). The CD spectrum of the DSE of I98A-CTL9 (green) was calculated by subtracting the population-weighted CD signal of the folded state (using the CD spectrum of wild type as the folded CD spectrum). The CD spectrum of I98A-CTL9 in 8 M urea (blue) is shown for comparison.

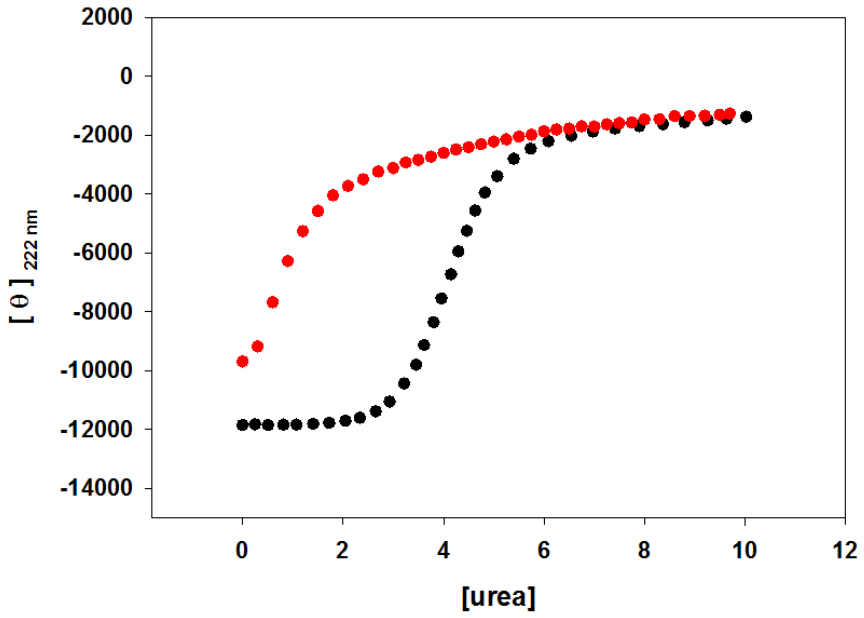


Figure 5-11. Urea unfolded of I98A-CTL9 (red) and wild type CTL9 (black) at 25°C, pH 6.0. The buffer consists of 10 mM DMG and 120 mM NaCl.

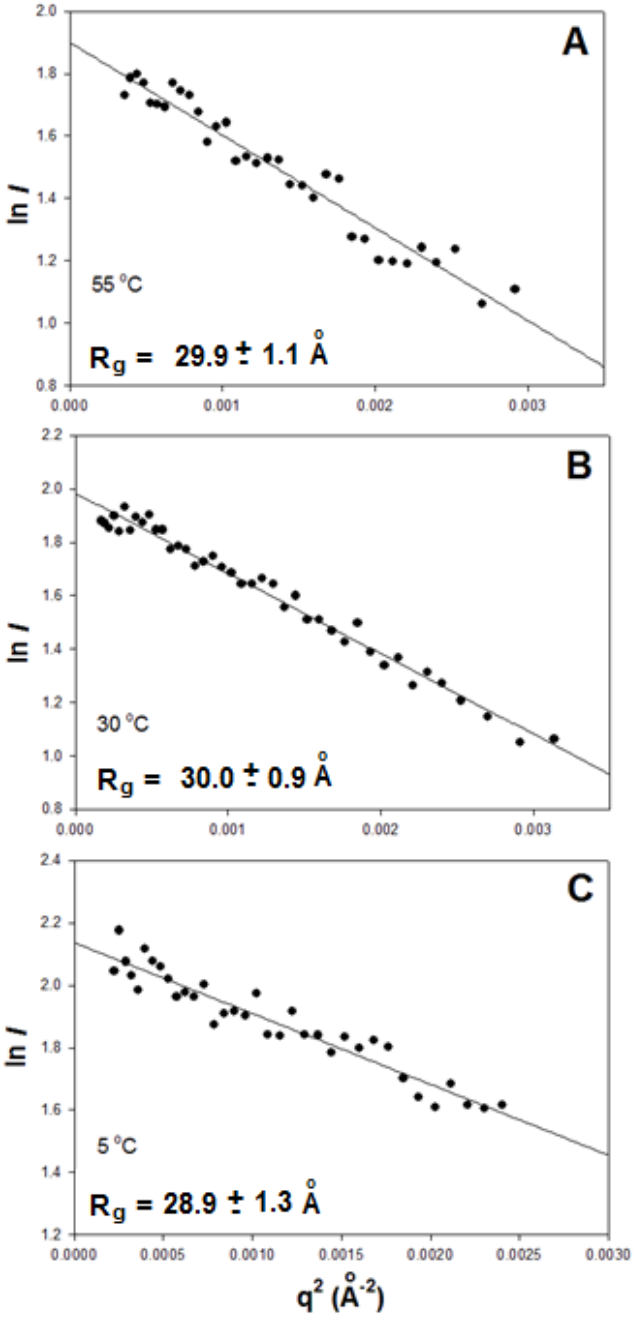


Figure 5-12. Guinier analysis of I98A-CTL9 in 6 M GdnHCl , pD 5.6, at (A) 55°C; (B) 30°C; and (C) 5°C. The R_g values obtained from the analysis agrees with the R_g value expected for an expanded random coil model. The buffer consists 10 mM DMG and 120 mM NaCl.

5.5. References:

[1] Dias, C. L. (2012) Unifying microscopic mechanism for pressure and cold denaturations of proteins, *Phys Rev Lett 109*, 048104.

[2] Graziano, G. (2010) On the molecular origin of cold denaturation of globular proteins, *Phys Chem Chem Phys 12*, 14245-14252.

[3] Dias, C. L., Ala-Nissila, T., Wong-ekkabut, J., Vattulainen, I., Grant, M., and Karttunen, M. (2010) The hydrophobic effect and its role in cold denaturation, *Cryobiology 60*, 91-99.

[4] Yoshidome, T., and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein, *Phys Rev E 79*, 030905.

[5] 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.

[6] Lopez, C. F., Darst, R. K., and Rossky, P. J. (2008) Mechanistic elements of protein cold denaturation, *J Phys Chem B 112*, 5961-5967.

[7] Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I., and Grant, M. (2008) Microscopic mechanism for cold denaturation, *Phys Rev Lett 100*, 118101.

[8] Tsai, C. J., Maizel, J. V., Jr., and Nussinov, R. (2002) The hydrophobic effect: A new insight from cold denaturation and a two-state water structure, *Crit Rev Biochem Mol Biol 37*, 55-69.

[9] Franks, F. (1995) Protein destabilization at low temperatures, *Adv Protein Chem 46*, 105-139.

[10] Privalov, P. L., Griko Yu, V., Venyaminov, S., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin, *J Mol Biol 190*, 487-498.

[11] Huang, G. W. S., and Oas, T. G. (1996) Heat and cold denatured states of monomeric λ repressor are thermodynamically and conformationally equivalent, *Biochemistry 35*, 6173-6180.

[12] Wuttke, R., Hofmann, H., Nettels, D., Borgia, M. B., Mittal, J., Best, R. B., and Schuler, B. (2014) Temperature-dependent solvation modulates the dimensions of disordered proteins, *Proc Natl Acad Sci USA 111*, 5213-5218.

[13] Aznauryan, M., Nettels, D., Holla, A., Hofmann, H., and Schuler, B. (2013) Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins, *J Am Chem Soc 135*, 14040-14043.

[14] Adrover, M., Martorell, G., Martin, S. R., Urosev, D., Konarev, P. V., Svergun, D. I., Daura, X., Temussi, P., and Pastore, A. (2012) The role of hydration in protein stability: Comparison of the cold and heat unfolded states of Yfh1, *J Mol Biol 417*, 413-424.

[15] 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.

[16] 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.

[17] Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyama, S., and Akasaka, K. (2006) Cold denaturation of ubiquitin at high pressure, *Magn Reson Chem 44 Spec No*, S108-113.

[18] 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.

[19] 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, *Nature Struct Mol Biol 11*, 352-357.

[20] Adrover, M., Esposito, V., Martorell, G., Pastore, A., and Temussi, P. A. (2010) Understanding cold denaturation: The case study of Yfh1, *J Am Chem Soc 132*, 16240-16246.

[21] Martin, S. R., Esposito, V., De Los Rios, P., 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.

[22] 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.

[23] Nettels, D., Muller-Spath, S., Kuster, F., Hofmann, H., Haenni, D., Ruegger, S., Reymond, L., Hoffmann, A., Kubelka, J., Heinz, B., Gast, K., Best, R. B., and Schuler, B. (2009) Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins, *Proc Natl Acad Sci USA 106*, 20740-20745.

[24] Luan, B. W., Shan, B., Baiz, C., Tokmakoff, A., and Raleigh, D. P. (2013) Cooperative cold denaturation: The case of the C-Terminal domain of ribosomal protein L9, *Biochemistry 52*, 2402-2409.

[25] Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2007) 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.

[26] Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318*, 571-582.

[27] Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure, *J Am Chem Soc 132*, 4669-4677.

[28] 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.

[29] Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions, *Protein Sci 3*, 843-852.

[30] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36*, 1277-1282.

[31] Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays*, Wiley, New York.

[32] Bernhard, S. A. (1956) Ionization constants and heats of tris(Hydroxymethyl)aminomethane and phosphate buffers, *J Biol Chem 218*, 961-969.

[33] 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., Hasan, 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.

[34] Kjaergaard, M., Norholm, A. B., Hendus-Altenburger, R., Pedersen, S. F., Poulsen, F. M., and Kragelund, B. B. (2010) Temperature-dependent structural changes in intrinsically disordered proteins: Formation of α-helices or loss of polyproline II?, *Protein Sci 19*, 1555-1564.

[35] Sun, S. T., Nishio, I., Swislow, G., and Tanaka, T. (1980) The coil-globule transition: Radius of gyration of polystyrene in cyclohexane, *J Chem Phys 73*, 5971-5975.

[36] Sanchez, I. C. (1979) Phase-transition behavior of the isolated polymer-chain, *Macromol 12*, 980-988.

[37] Uversky, V. N. (2009) Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding, *Protein J 28*, 305-325.

[38] Voelz, V. A., Singh, V. R., Wedemeyer, W. J., Lapidus, L. J., and Pande, V. S. (2010) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment, *J Am Chem Soc 132*, 4702-4709.

[39] Wang, Z., Plaxco, K. W., and Makarov, D. E. (2007) Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins, *Biopolymers 86*, 321-328.

[40] Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc Natl Acad Sci USA 101*, 12497-12502.

Chapter 6

Conclusions and Perspectives from This Thesis

In this dissertation, the properties of the CTL9 denatured state ensemble (DSE) under a wide range of conditions were studied. SAXS was used to obtain the overall dimension, as judged by R_g, while NMR was used to provide atomic level residue-specific information.

It is widely thought that theories adapted from polymer physics[1] are well suited to describe the protein denatured state conformational ensemble populated in strongly denaturing conditions.[2] In particular, the scaling behavior of chain size, as defined by the average radius of gyration (R_g) with chain length (N, the number of residues), can be used to quantify the dimension of the DSE. This power law has the form of $R_g = R_0 N^v$, as shown by Flory.[1] The constant R_0 is a function of the polymer persistence length and v is a scaling factor which depends on the solvent quality. Depending on the solution conditions, R_0 and v have different values. If v has a value of 0.6, it indicates that the chain is in a good solvent and expands to make favorable contacts with the surrounding solvent. Whereas a value of 0.34 for v is indicative of the chain forming a compact globule, minimizing the contacts between the chain and the surrounding poor solvent. A good solvent is formally defined as a solvent in which a chain makes more favorable interactions with the surrounding solvent, i.e. the chain-solvent interactions are more preferable to the chain-chain interactions.

The hypothesis that highly denatured proteins induced by 8 M urea or 6 M GdnHCl are highly expanded and behave like chains in good solvents[3] is supported by SAXS measurements of R_g as a function of N for 28 proteins, with chain lengths varying from 8 to 549 residues.[4] In that work, the measured R_g is consistent with the scaling factor of ~ 0.6.

In a "perfect solvent", the DSE can be modeled by ignoring all interactions except for steric repulsive interactions of the excluded volume (EV) type, and a $N^{0.59}$ law is obeyed for both short and long chains in the EV limit. The perfect solvent condition can be defined as the scenario where chain-solvent interactions counter-balance all non-EV intra-chain interactions, as a result, the limit of the perfect solvent is also termed as the EV limit.[5] Both the R_0 and ν derived from the experimental studies of R_g were examined by Tran et al. These workers generated equilibrium ensembles following the EV limit for the 28 protein sequences that were studied by Kohn et al.[6] The values of ν and R_0 calculated by Tran et al. showed statistically significant agreement with the estimated values from the SAXS data.[4, 6] Thus, harshly denaturing conditions (as opposed to mildly denaturing conditions) can be deemed as close to a "perfect solvent" rather than just a good solvent as judged by SAXS. Those findings suggest that EV protein ensembles are close to the actual protein DSE populated in high concentrations of chemical denaturant, at least as judged by R_g. However, it is not clear if the distribution of contacts within the DSE are consistent with the EV limit. It is also not clear if DSEs that have same R_g, but are populated under different conditions have the same distribution of the contacts.

While R_g is a useful parameter to deduce if the protein is unfolded, long-range transient contacts can exist within the DSE. Several examples have been reported for the coexistence of an expanded R_g that follows random coil scaling laws and residual structure (long-range contacts) within the DSE.[7-10] The reconciliation of the random-coil parameters and the ordered residual structure formed within the DSE can be explained by the fact that R_g of a denatured ensemble is not sensitive to the detailed structure of the unfolded chain.[11] The work in this thesis provides additionally direct evidence of this behavior as demonstrated by studies of the acid-induced and low-pH urea-induced DSE of CTL9. Based on SAXS data, the dimension of both the above mentioned DSE's follow the scaling law behavior expected for a highly unfolded coil. However, it was shown that the pattern of long-range interactions formed within the acid-induced DSE clearly deviates from what one expects for the EV model.

The acid-induced DSE has been studied for several proteins by SAXS. Just like the CTL9 DSE, the acid induced DSE of cytochrome c and staphylococcal nuclease (Snase) are as expanded as the DSE induced by urea or GdnHCl.[12, 13] In contrast, the apomyoglobin acid-induced DSE is more compact than the DSE generated by chemical denaturant.[14] In those studies, the pattern of the long-range contacts formed in those acid-DSEs were not reported. Charge repulsion induced by the high charge state at low pH is the most intuitive explanation for the expanded acid-DSE, however there are examples where a significant change in charge results in only a 10% increase in R_g.[15] For CTL9, the acid-induced DSE and the urea-induced DSE at low pH values, have very similar values

of R_g. These results of this thesis provide more information for studying the properties of the DSE induced by different means.

CTL9 was shown to be fully populate the DSE as judged by CD at pH 2.0 no urea and at pH 2.5 with 8 M urea.[16] The experimental R_g values of both the low-pH urea- and acid-induced DSE are similar, and they are consistent with the R_g values predicted for expanded polymers in a good solvent according to the scaling laws. However, as shown by the NMR secondary structure propensity (SSP) analysis, the propensity to form secondary structure in the two DSEs are different: There is a significant propensity for the acid DSE to populate the helical structure, whereas the urea DSE has no such propensity. The acid-induced DSE of CTL9 has average SSP scores of 0.28 for helix-1 and 0.32 for helix-2 (a SSP score of 1 indicates fully formed α-helix while -1 means fully formed β-sheet), whereas the low-pH urea-induced DSE of CTL9 has an average of SSP score close to 0. NMR-PRE measurements are sensitive to transient long-range intramolecular interactions within the DSE and to the transiently formed compact structures, at a population of as low as 1%. Individual contacts can be detected at even lower levels. The PRE studies confirmed that the acid-induced DSE of CTL9 has extensive long-range contacts formed, i.e. it deviates from the EV model. While the addition of urea (8 M urea at pH 2.5) generates a DSE that behaves as one would expected for an EV model. These results confirm that parameters such as R_g are not sensitive to transient long-range contacts or low populations of compact structures. In order to describe the protein DSE more accurately, atomic level information provided by

NMR-PRE and/or by single molecule FRET are needed, both methods containing a distance dependence of r^{-6}, adding more information in addition to R_g.

Protein denaturation caused by low and high temperatures follow different mechanisms. Cold denaturation is mainly driven by decreased hydrophobic interactions within the protein, together with stronger hydrogen bonds forming with water at low temperatures. On the other hand, in addition to the temperature dependent hydrophobic effects, protein heat or thermal denaturation is caused by the increased entropy of the polypeptide chain, since the entropic contribution to ΔG^o is $-T\Delta S^o$ and, all else being equal, will become more important at a higher temperature.[17-19]

Generally speaking, thermal induced unfolded proteins are more compact than the equivalent chemical denatured proteins. It has been suggested that the thermal denaturation is closer to the theta conditions, where the solvent is poor enough to cancel the expansion effects caused by excluded volume, as a result, the chain contracts. Correspondingly a more compact thermal DSE is formed.[11] Compaction of 10-25% has been reported for the thermally unfolded DSE relative to the denaturant induced DSE for cytochrome *c*[20], neocarzinostatin[21], and subtilisin inhibitor[22]. Under reducing conditions, thermally denatured hen egg white lysozyme and ribonuclease A have R_g values similar to the denaturant induced DSE, respectively.[23, 24] Once thermal unfolding is achieved, the R_g of the heat induced DSE has been found to be independent of any subsequent increase in temperature for some proteins.[24, 25] But this is still controversial: Schuler et al. reported a compaction for thermally induced DSE from smFRET measurements at

elevated temperatures.[26] This topic has been addressed in this thesis, CTL9 has a relative compact heat induced DSE which shows no temperature dependence in the temperature range studied in this thesis. This result is in agreement with other SAXS studies, but disagrees with the results of Schuler et al.

The work in this thesis also addresses the properties of the cold induced DSE. Previous studies of some proteins cold denatured in the presence of chemical denaturant imply that the cold induced DSE is highly unstructured based on spectroscopic data, but the ensembles are more compact than those induced by chemical denaturants at high concentrations, as judged by the R_g. For example, the cold-denatured β–lactoglobulin is compact with an R_g that is 20% larger than its folded state but 35% smaller than the GdnHCl denatured state at 0 °C; in addition, the Kratky plot suggests that the geometry of the ensemble lies between an extended coil and a compact globule structure.[27] General conclusions about the properties of the protein cold denatured state under near native conditions have yet to be achieved, due to the limited number of studies of protein cold denaturation that did not involve lowering the pH or adding denaturant. My work with the I98A mutant of CTL9 helps to bridge this gap. The cold denatured DSE of I98A-CTL9 has a temperature-dependent expansion. The DSE expands as the temperature is lowered, and the R_g shows an expansion of ~ 80% compared with its folded state at 5 °C, compared to an expansion of ~30% at 25 °C.

I98A-CTL9 also enables the direct comparison of the heat and cold DSE for the same protein, revealing different behaviors between the two DSE's. A relatively compact heat

induced CTL9 DSE (an expansion of ~ 50% compared with its folded state), with no temperature dependence is observed, in contrast to the expansion of the cold denatured DSE. The heat and cold DSE also differ in terms of the secondary structure content. The cold DSE appears to have more helical structure as judged by NMR, compared to the thermally induced DSE. It may be the case that the cold denatured DSE forms slightly more ordered helical content, due to the decrease of the entropy and water ordering around the protein. Whereas in the heat induced DSE, the entropic contribution at higher temperature makes the protein less likely to have a significant propensity to form secondary structure. The fact that the compact heat DSE has less helical content may seem counterintuitive, but it consistent with the fact that there is no strong correlation between secondary structure and the compactness for IDPs.[28]

The data presented in this thesis advances the field by exploring the understudied process of cold denaturation, adding a crucial piece of information to help rationalize the expanded dimensions and residual structural of protein DSE under certain conditions. Furthermore, a direct characterization and comparison of both the cold- and heat-induced DSE on the same protein reveals differences between the two ensembles.

In order to further understand the correlations, if any, among protein DSEs populated under different conditions, proposed future studies should include the examination of the acid- and urea-induced CTL9 DSE as a function of temperature. At room temperature, both the acid- and low-pH urea-induced DSE can be deemed as having random coil R_g values. By investigating those two DSEs at different temperatures, another benchmark for

the random coils that takes temperature dependent factors into account can be established. Furthermore, any temperature dependence observed for the acid- and low-pH urea-induced DSE may help explain the difference between the heat and cold induced I98A-CTL9 DSE populated in buffer. This is of interest since the most relevant DSE for folding is the ensemble populated at native pH in the absence of denaturant.

References:

[1] Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Itchaca, NY.

[2] Chan, H. S., and Dill, K. A. (1991) Polymer principles in protein-structure and stability, *Annu Rev Biophys Bio 20*, 447-490.

[3] Tanford, C. (1968) Protein denaturation, *Adv Protein Chem 23*, 121-282.

[4] 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., Hasan, 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 Acade Sci USA101*, 12491-12496.

[5] Rubenstein, M. (2003) *Polymer Physics*, Oxford University Press, New York.

[6] Tran, H. T., and Pappu, R. V. (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions, *Biophys J 91*, 1868-1886.

[7] Meng, W., Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2013) The denatured state ensemble contains significant local and long-range structure under native conditions: analysis of the N-terminal domain of ribosomal protein L9, *Biochemistry 52*, 2662-2671.

[8] Voelz, V. A., Singh, V. R., Wedemeyer, W. J., Lapidus, L. J., and Pande, V. S. (2010) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment, *J Am Chem Soc 132*, 4702-4709.

[9] Wang, Z., Plaxco, K. W., and Makarov, D. E. (2007) Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins, *Biopolymers 86*, 321-328.

[10] Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc Natl Acade Sci USA 101*, 12497-12502.

[11] Millett, 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 62*, 241-262.

[12] Uversky, V. N., Karnoup, A. S., Segel, D. J., Seshadri, S., Doniach, S., and Fink, A. L. (1998) Anion-induced folding of Staphylococcal nuclease: Characterization of multiple equilibrium partially folded intermediates, *J Mol Biol 278*, 879-894.

[13] Kamatari, Y. O., Konno, T., Kataoka, M., and Akasaka, K. (1996) The methanol-induced globular and expanded denatured states of cytochrome c: A study by CD fluorescence, NMR and small-angle X-ray scattering, *J Mol Biol 259*, 512-523.

[14] Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, Y. (1995) Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray-scattering, *J Mol Biol 249*, 215-228.

[15] Millet, I. S., Townsley, L. E., Chiti, F., Doniach, S., and Plaxco, K. W. (2002) Equilibrium collapse and the kinetic 'foldability' of proteins, *Biochemistry 41*, 321-325.

[16] Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9, *J Mol Biol 349*, 839-846.

[17] Graziano, G. (2014) On the mechanism of cold denaturation, *Phys Chem Chem Phys 16*, 21755-21767.

[18] Yoshidome, T., and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein, *Phys Rev E 79*, 030905.

[19] Nettels, D., Muller-Spath, S., Kuster, F., Hofmann, H., Haenni, D., Ruegger, S., Reymond, L., Hoffmann, A., Kubelka, J., Heinz, B., Gast, K., Best, R. B., and

Schuler, B. (2009) Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins, *Proc Natl Acade Sci USA 106*, 20740-20745.

[20] Hagihara, Y., Hoshino, M., Hamada, D., Kataoka, M., and Goto, Y. (1998) Chain-like conformation of heat-denatured ribonuclease A and cytochrome c as evidenced by solution X-ray scattering, *Fold Des 3*, 195-201.

[21] Perez, J., Vachette, P., Russo, D., Desmadril, M., and Durand, D. (2001) Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering, *J Mol Biol 308*, 721-743.

[22] Konno, T., Kamatari, Y. O., Kataoka, M., and Akasaka, K. (1997) Urea-induced conformational changes in cold- and heat-denatured states of a protein, Streptomyces subtilisin inhibitor, *Protein Sci 6*, 2242-2249.

[23] Arai, S., and Hirai, M. (1999) Reversibility and hierarchy of thermal transition of hen egg-white lysozyme studied by small-angle X-ray scattering, *Biophys J 76*, 2192-2197.

[24] Jacob, J., Dothager, R. S., Thiyagarajan, P., and Sosnick, T. R. (2007) Fully reduced ribonuclease a does not expand at high denaturant concentration or temperature, *J Mol Biol 367*, 609-615.

[25] Panick, G., Malessa, R., Winter, R., Rapp, G., Frye, K. J., and Royer, C. A. (1998) Structural characterization of the pressure-denatured state and unfolding/refolding kinetics of staphylococcal nuclease by synchrotron small-angle X-ray scattering and Fourier-transform infrared spectroscopy, *J Mol Biol 275*, 389-402.

[26] Aznauryan, M., Nettels, D., Holla, A., Hofmann, H., and Schuler, B. (2013) Single-molecule spectroscopy of cold denaturation and the temperature-induced collapse of unfolded proteins, *J Am Chem Soc 135*, 14040-14043.

[27] Katou, H., Hoshino, M., Kamikubo, H., Batt, C. A., and Goto, Y. (2001) Native-like beta-hairpin retained in the cold-denatured state of bovine beta-lactoglobulin, *J Mol Biol 310*, 471-484.

[28] Marsh, J. A., and Forman-Kay, J. D. (2010) Sequence determinants of compaction in intrinsically disordered proteins, *Biophys J 98*, 2383-2390.

Complete List of References:

Adrover, M., Esposito, V., Martorell, G., Pastore, A., and Temussi, P. A. (2010) Understanding cold denaturation: The case study of Yfh1, *J Am Chem Soc 132,* 16240-16246.

Adrover, M., Martorell, G., Martin, S. R., Urosev, D., Konarev, P. V., Svergun, D. I., Daura, X., Temussi, P., and Pastore, A. (2012) The role of hydration in protein stability: Comparison of the cold and heat unfolded States of Yfh1, *J Mol Biol 417,* 413-424.

Allison, J. R., Varnai, P., Dobson, C. M., and Vendruscolo, M. (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements, *J Am Chem Soc 131,* 18314-18326.

Anfinsen, C. B. (1972) The formation and stabilization of protein structure, *Biochem J 128,* 737-749.

Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181,* 223-230.

Anil, B., Song, B. B., Tang, Y. F., and Raleigh, D. P. (2004) Exploiting the right side of the ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability, *J Am Chem Soc 126,* 13194-13195.

Auton, M., Holthauzen, L. M. F., and Bolen, D. W. (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation, *Proc Natl Acad Sci USA 104,* 15317-15322.

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.

Baiz, C. R., Peng, C. S., Reppert, M. E., Jones, K. C., and Tokmakoff, A. (2012) Coherent two-dimensional infrared spectroscopy: Quantitative analysis of protein secondary structure in solution, *Analyst 137,* 1793-1799.

Baker, D. (2000) A surprising simplicity to protein folding, *Nature 405,* 39-42.

Baldwin, R. L. (2002) A new perspective on unfolded proteins, *Adv Protein Chem 62,* 361-367.

Barrick, D. (2009) What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding?, *Phys Biol 6,* 015001.

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.

Bernado, P., and Blackledge, M. (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering, *Biophys J* 97, 2839-2845.

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.

Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering, *J Am Chem Soc 129,* 5656-5664.

Bolin, K. A., Pitkeathly, M., Miranker, A., Smith, L. J., and Dobson, C. M. (1996) Insight into a random coil conformation and an isolated helix: Structural and dynamical characterisation of the C-helix peptide from hen lysozyme, *J Mol Biol 261,* 443-453.

Bowler, B. E. (2007) Thermodynamics of protein denatured states, *Mol BioSys 3,* 88-99.

Bradley, P., Misura, K. M., and Baker, D. (2005) Toward high-resolution *de novo* structure prediction for small proteins, *Science 309,* 1868-1871.

Brandts, J. F. (1964) Thermodynamics of protein denaturation .2. Model of reversible denaturation and interpretations regarding stability of chymotrypsinogen, *J Am Chem Soc 86,* 4302-4302.

Brandts, J. F. (1964) Thermodynamics of protein denaturation .I. Denaturation of chymotrypsinogen, *J Am Chem Soc 86,* 4291-4292.

Brimacombe, R., Gornicki, P., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, 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.

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.

Canchi, D. R., and Garcia, A. E. (2011) Backbone and side-chain contributions in protein denaturation by urea, *Biophys J 100,* 1526-1533.

Candotti, M., Esteban-Martin, S., Salvatella, X., and Orozco, M. (2013) Toward an atomistic description of the urea-denatured state of proteins, *Proc Natl Acad Sci USA 110,* 5933-5938.

Chan, H. S., and Dill, K. A. (1998) Protein folding in the landscape perspective: Chevron plots and non-Arrhenius kinetics, *Proteins 30,* 2-33.

Chan, H. S., Shimizu, S., and Kaya, H. (2004) Cooperativity principles in protein folding, *Methods Enzymol 380,* 350-379.

Chen, B. L., and Schellman, J. A. (1989) Low-temperature unfolding of a mutant of phage-T4 lysozyme .1. Equilibrium studies, *Biochemistry 28,* 685-691.

Chen, B. L., Baase, W. A., and Schellman, J. A. (1989) Low-temperature unfolding of a mutant of phage-T4 lysozyme .2. Kinetic investigations, *Biochemistry 28,* 691-699.

Chen, L. L., Wildegger, G., Kiefhaber, T., Hodgson, K. O., and Doniach, S. (1998) Kinetics of lysozyme refolding: Structural characterization of a non-specifically collapsed state using time-resolved X-ray scattering, *J Mol Biol 276,* 225-237.

Cheung, M. S., Klimov, D., and Thirumalai, D. (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins, *Proc Natl Acad Sci USA 102,* 4753-4758.

Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease, *Annu Rev Biochem 75,* 333-366.

Cho, J. H., and Raleigh, D. P. (2006) Denatured state effects and the origin of nonclassical ϕ values in protein folding, *J Am Chem Soc 128,* 16492-16493.

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.

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.

Choy, W. Y., Mulder, F. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and Kay, L. E. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques, *J Mol Biol 316,* 101-112.

Clore, G. M., and Gronenborn, A. M. (1998) Determining the structures of large proteins and protein complexes by NMR, *Trends Biotechnol 16,* 22-34.

Clore, G. M., and Gronenborn, A. M. (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa, *Curr Opin Chem Biol 2,* 564-570.

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.

Daggett, V., and Fersht, A. (2003) The present view of the mechanism of protein folding, *Nat Rev Mol Cell Biol 4,* 497-502.

Dalby, P. A., Oliveberg, M., and Fersht, A. R. (1998) Folding intermediates of wild-type and mutants of barnase. I. Use of ϕ-value analysis and m-values to probe the cooperative nature of the folding pre-equilibrium, *J Mol Biol 276,* 625-646.

Das, R. K., and Pappu, R. V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues, *Proc Natl Acad Sci USA 110,* 13392-13397.

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.

Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations, *J Am Chem Soc 127,* 476-477.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processingsystem based on Unix Pipes, *J Biomol NMR 6,* 277-293.

Dias, C. L. (2012) Unifying microscopic mechanism for pressure and cold denaturations of proteins, *Phys Rev Lett 109,* 048104.

Dias, C. L., Ala-Nissila, T., Karttunen, M., Vattulainen, I., and Grant, M. (2008) Microscopic mechanism for cold denaturation, *Phys Rev Lett 100,* 118101.

Dias, C. L., Ala-Nissila, T., Wong-ekkabut, J., Vattulainen, I., Grant, M., and Karttunen, M. (2010) The hydrophobic effect and its role in cold denaturation, *Cryobiology 60,* 91-99.

Dill, K. A. (1985) Theory for the folding and stability of globular proteins, *Biochemistry 24,* 1501-1509.

Dill, K. A., and MacCallum, J. L. (2012) The protein-folding problem, 50 years on, *Science 338,* 1042-1046.

Ding, F., Jha, R. K., and Dokholyan, N. V. (2005) Scaling behavior and structure of denatured proteins, *Structure 13,* 1047-1054.

Dobson, C. M., and Karplus, M. (1999) The fundamentals of protein folding: Bringing together theory and experiment, *Curr Opin Struct Biol 9,* 92-101.

Dobson, C. M., Šali, A., and Karplus, M. (1998) Protein Folding: A perspective from theory and experiment, *Angew Chem Int Ed 37,* 868-893.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *J Mol Graphics Modell 19,* 26-59.

Dyson, H. J., and Wright, P. E. (1998) Equilibrium NMR studies of unfolded and partially folded proteins, *Nat Struct Biol 5 Suppl,* 499-503.

Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR, *Chem Rev 104*, 3607-3622.

Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nature Rev Mol Cell Biol 6*, 197-208.

Dyson, H. J., Wright, P. E., and Scheraga, H. A. (2006) The role of hydrophobic interactions in initiation and propagation of protein folding, *Proc Natl Acad Sci USA 103*, 13057-13061.

Eisenberg, D., Nelson, R., Sawaya, M. R., Balbirnie, M., Sambashivan, S., Ivanova, M. I., Madsen, A. O., and Riekel, C. (2006) The structural biology of protein aggregation diseases: Fundamental questions and some answers, *Acc Chem Res 39,* 568-575.

Eliezer, D. (2007) Characterizing residual structure in disordered protein States using nuclear magnetic resonance, *Methods Mol Biol 350,* 49-67.

England, J. L., and Haran, G. (2011) Role of solvation effects in protein denaturation: From thermodynamics to single molecules and back, *Annu Rev Phys Chem 62*, 257-277.

Englander, S. W., Mayne, L., and Krishna, M. M. (2007) Protein folding and misfolding: Mechanism and principles, *Q Rev Biophys 40*, 287-326.

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.

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.

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.

Feller, G. (2010) Protein stability and enzyme activity at extreme biological temperatures, *J Phys Condens Mat 22*, 323101-323117.

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.

Fersht, A. R. (1997) Nucleation mechanisms in protein folding, *Curr Opin Struct Biol 7*, 3-9.

Fersht, A. R., and Daggett, V. (2002) Protein folding and unfolding at atomic resolution, *Cell 108*, 573-582.

Fersht, A. R., Itzhaki, L. S., elMasry, N. F., Matthews, J. M., and Otzen, D. E. (1994) Single versus parallel pathways of protein folding and fractional formation of structure in the transition state, *Proc Natl Acad Sci USA 91,* 10426-10429.

Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding, *J Mol Biol 224,* 771-782.

Fieber, W., Kristjansdottir, S., and Poulsen, F. M. (2004) Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings, *J Mol Biol 339,* 1191-1199.

Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., and Dobson, C. M. (1996) Toward a description of the conformations of denatured states of proteins. Comparison of a random coil model with NMR measurements, *J Phys Chem 100,* 2661-2666.

Fitzkee, N. C., and Rose, G. D. (2004) Reassessing random-coil statistics in unfolded proteins, *Proc Natl Acad Sci USA 101,* 12497-12502.

Flory, P. J. (1953) *Principles of Polymer Chemistry,* Cornell University Press, Itchaca, NY.

Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules,* Wiley, New York.

Franks, F. (1995) Protein destabilization at low temperatures, *Adv Protein Chem 46,* 105-139.

Franzmann, M., Otzen, D., and Wimmer, R. (2009) Quantitative use of parramagnetic relaxation enhancements for determining orientations and insertion depths of peptides in micelles, *Chem Bio Chem 10,* 2339-2347.

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.

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.

Graziano, G. (2010) On the molecular origin of cold denaturation of globular proteins, *Phys Chem Chem Phys 12,* 14245-14252.

Griko, Y. V., Privalov, P. L., Sturtevant, J. M., and Venyaminov, S. Y. (1988) Cold denaturation of staphylococcal nuclease, *Proc Natl Acad Sci USA 85,* 3343-3347.

Guinier, A., and Fournet, G. (1955) *Small Angle Scattering of X-Rays,* Wiley, New York.

Hart, T., Hosszu, L. L., 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.

Hayley, M., Chevaldina, T., and Heeley, D. H. (2011) Cold adaptation of tropomyosin, *Biochemistry 50,* 6559-6566.

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.

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 bi-lobed RNA-binding protein, *EMBO J 13,* 205-212.

Horng, J. C., Cho, J. H., and Raleigh, D. P. (2005) Analysis of the pH-dependent folding and stability of histidine point mutants allows characterization of the denatured state and transition state for protein folding, *J Mol Biol 345,* 163-173.

Horwich, A. L., Farr, G. W., and Fenton, W. A. (2006) GroEL-GroES-mediated protein folding, *Chem Rev 106,* 1917-1930.

Horwich, A. L., Weber-Ban, E. U., and Finley, D. (1999) Chaperone rings in protein folding and degradation, *Proc Natl Acad Sci USA 96,* 11033-11040.

Hoshino, M., Hagihara, Y., Hamada, D., Kataoka, M., and Goto, Y. (1997) Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering, *FEBS Lett 416,* 72-76.

Huang, J. R., Gabel, F., Jensen, M. R., Grzesiek, S., and Blackledge, M. (2012) Sequence-specific mapping of the interaction between urea and unfolded ubiquitin from ensemble analysis of NMR and small angle scattering data, *J Am Chem Soc 134,* 4429-4436.

Iwahara, J., and Clore, G. M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR, *Nature 440,* 1227-1230.

Jackson, S. E. (1998) How do small single-domain proteins fold?, *Fold Des 3,* R81-R91.

Jacob, J., Dothager, R. S., Thiyagarajan, P., and Sosnick, T. R. (2007) Fully reduced ribonuclease a does not expand at high denaturant concentration or temperature, *J Mol Biol 367,* 609-615.

Jacob, J., Krantz, B., Dothager, R. S., Thiyagarajan, P., and Sosnick, T. R. (2004) Early collapse is not an obligate step in protein folding, *J Mol Biol 338,* 369-382.

Jacobsen, C. F., and Christensen, L. K. (1948) Influence of temperature on the urea denaturation of β-lactoglobulin, *Nature 161,* 30-31.

Jahn, T. R., and Radford, S. E. (2008) Folding versus aggregation: Polypeptide conformations on competing pathways, *Arch Biochem Biophys 469,* 100-117.

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.

Johansson, J., Gudmundsson, G. H., Rottenberg, M. E., Berndt, K. D., and Agerberth, B. (1998) Conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37, *J Biol Chem 273*, 3718-3724.

Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol 278*, 313-352.

Jones, J., Wilkins, D., Smith, L., and Dobson, C. (1997) Characterisation of protein unfolding by NMR diffusion measurements, *J Biomol NMR 10*, 199-203.

Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *J Phys Chem B 105*, 6474-6487.

Karplus, M., and Weaver, D. L. (1994) Protein folding dynamics: The diffusion-collision model and experimental data, *Protein Sci 3*, 650-668.

Kaya, H., and Chan, H. S. (2000) Polymer principles of protein calorimetric two-state cooperativity, *Proteins 40*, 637-661.

Kaya, H., and Chan, H. S. (2003) Simple two-state protein folding kinetics requires near-levinthal thermodynamic cooperativity, *Proteins 52*, 510-523.

Kim, H. Y., Cho, M. K., Riedel, D., Fernandez, C. O., and Zweckstetter, M. (2008) Dissociation of amyloid fibrils of α-synuclein in supercooled water, *Angew Chem Int Ed Engl 47*, 5046-5048.

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.

Kimura, T., Akiyama, S., Uzawa, T., Ishimori, K., Morishima, I., Fujisawa, T., and Takahashi, S. (2005) Specifically collapsed intermediate in the early stage of the folding of ribonuclease A, *J Mol Biol 350*, 349-362.

Kimura, T., Uzawa, T., Ishimori, K., Morishima, I., Takahashi, S., Konno, T., Akiyama, S., and Fujisawa, T. (2005) Specific collapse followed by slow hydrogen-bond formation of β-sheet in the folding of single-chain monellin, *Proc Natl Acad Sci USA 102*, 2748-2753.

Kitahara, R., Okuno, A., Kato, M., Taniguchi, Y., Yokoyama, S., and Akasaka, K. (2006) Cold denaturation of ubiquitin at high pressure, *Magn Reson Chem 44*, S108-S113.

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.

Klimov, D. K., Straub, J. E., and Thirumalai, D. (2004) Aqueous urea solution destabilizes A β(16-22) oligomers, *Proc Natl Acad Sci USA 101,* 14760-14765.

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., Hasan, 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.

Kolhe, P., and Badkar, A. (2011) Protein and solute distribution in drug substance containers during frozen storage and post-thawing: A tool to understand and define freezing-thawing parameters in biotechnology process development, *Biotechnol Progr 27,* 494-504.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36,* 1277-1282.

Konuma, T., Kimura, T., Matsumoto, S., Got, Y., Fujisawa, T., Fersht, A. R., and Takahashi, S. (2011) Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase, *J Mol Biol 405,* 1284-1294.

Kosen, P. A., Scheek, R. M., Naderi, H., Basus, V. J., Manogaran, S., Schmidt, P. G., Oppenheimer, N. J., and Kuntz, I. D. (1986) Two-dimensional ^{1}H NMR of three spin-labeled derivatives of bovine pancreatic trypsin inhibitor, *Biochemistry 25,* 2356-2364.

Kresheck, G. C., Schneider, H., and Scheraga, H. A. (1965) The effect of D_2O on the thermal stability of proteins. Thermodynamic parameters for the transfer of model compounds from H_2O to D_2O, *J Phys Chem 69,* 3132-3144.

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.

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.

Kuhlman, B., Luisi, D. L., Young, P., and Raleigh, D. P. (1999) pKa values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions, *Biochemistry 38,* 4896-4903.

Lazar, K. L., Patapoff, T. W., and Sharma, V. K. (2010) Cold denaturation of monoclonal antibodies, *mAbs 2,* 42-52.

Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009) Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP, *Biochemistry 48*, 2115-2124.

Levintha.C. (1968) Are there pathways for protein folding, *J Chim Phys 65*, 44-45.

Li, Y., Gupta, R., Cho, J. H., and Raleigh, D. P. (2007) 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.

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.

Li, Y., Picart, F., and Raleigh, D. P. (2005) Direct characterization of the folded, unfolded and urea-denatured states of the C-terminal domain of the ribosomal protein L9, *J Mol Biol 349*, 839-846.

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.

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.

Lim, W. K., Rosgen, J., and Englander, S. W. (2009) Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, *Proc Natl Acad Sci USA 106*, 2595-2600.

Lindorff-Larsen, 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.

Lopez, C. F., Darst, R. K., and Rossky, P. J. (2008) Mechanistic elements of protein cold denaturation, *J Phys Chem B 112*, 5961-5967.

Luan, B. W., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2014) Denatured state ensembles with the same radii of gyration can form significantly different long-range contacts, *Biochemistry 53*, 39-47.

Luan, B., Shan, B., Baiz, C., Tokmakoff, A., and Raleigh, D. P. (2013) Cooperative cold denaturation: The case of the C-terminal domain of ribosomal protein L9, *Biochemistry 52*, 2402-2409.

Lumry, R., and Biltonen, R. (1966) Validity of the "two-state" hypothesis for conformational transitions of proteins, *Biopolymers 4*, 917-944.

Lynn, A., Chandra, S., Malhotra, P., and Chauhan, V. S. (1999) Heme binding and polymerization by Plasmodium falciparum histidine rich protein II: Influence of pH on activity and conformation, *FEBS Lett 459,* 267-271.

Ma, Y., and Hendershot, L. M. (2004) The role of the unfolded protein response in tumour development: Friend or foe?, *Nat Rev Cancer 4,* 966-977.

Makhatadze, G. I., Clore, G. M., and Gronenborn, A. M. (1995) Solvent isotope effect and protein stability, *Nat Struct Biol 2,* 852-855.

Manning, M. C., and Woody, R. W. (1991) Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors, *Biopolymers 31,* 569-586.

Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L., and Pappu, R. V. (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins, *Proc Natl Acad Sci USA 107,* 8183-8188.

Marley, J., Lu, M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J Biomol NMR 20,* 71-75.

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.

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.

McCarney, E. R., Kohn, J. E., and Plaxco, K. W. (2005) Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins, *Crit Rev Biochem Mol Biol 40,* 181-189.

Meng, W., Luan, B., Lyle, N., Pappu, R. V., and Raleigh, D. P. (2013) The denatured state ensemble contains significant local and long-range structure under native conditions: Analysis of the N-terminal domain of ribosomal protein L9, *Biochemistry 52,* 2662-2671.

Meng, W., Lyle, N., Luan, B., Raleigh, D. P., and Pappu, R. V. (2013) Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure, *Proc Natl Acad Sci USA 110,* 2123-2128.

Millet, I. S., Townsley, L. E., Chiti, F., Doniach, S., and Plaxco, K. W. (2002) Equilibrium collapse and the kinetic 'foldability' of proteins, *Biochemistry 41,* 321-325.

Millett, 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 62,* 241-262.

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.

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.

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.

Moult, J. (2005) A decade of CASP: Progress, bottlenecks and prognosis in protein structure prediction, *Curr Opin Struct Biol 15,* 285-289.

Mountain, R. D., and Thirumalai, D. (2003) Molecular dynamics simulations of end-to-end contact formation in hydrocarbon chains in water and aqueous urea solution, *J Am Chem Soc 125,* 1950-1957.

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.

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.

O'Brien, E. P., Dima, R. I., Brooks, B., and Thirumalai, D. (2007) Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions: Lessons for protein denaturation mechanism, *J Am Chem Soc 129,* 7346-7353.

Ohnishi, S., and Shortle, D. (2003) Observation of residual dipolar couplings in short peptides, *Proteins 50,* 546-551.

Onuchic, J. N., and Wolynes, P. G. (2004) Theory of protein folding, *Curr Opin Struct Biol 14,* 70-75.

Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol 131,* 266-280.

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.

Pace, N. C., and Tanford, C. (1968) Thermodynamics of unfolding of β-lactoglobulin a in aqueous urea solutions between 5 and 55 degrees, *Biochemistry 7,* 198-206.

Paroutis, P., Touret, N., and Grinstein, S. (2004) The pH of the secretory pathway: measurement, determinants, and regulation, *Physiology 19,* 207-215.

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,* 5374-5374.

Pavitt, G. D., and Ron, D. (2012) New insights into translational regulation in the endoplasmic reticulum unfolded protein response, *Cold Spring Harbor Perspect Biol 4*, DOI: 10.1101/cshperspect.a012278.

Perez, Y., Gairi, M., Pons, M., and Bernado, P. (2009) Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: Insights into the role of phosphorylation of the unique domain, *J Mol Biol 391,* 136-148.

Plaxco, K. W., Millett, I. S., Segel, D. J., Doniach, S., and Baker, D. (1999) Chain collapse can occur concomitantly with the rate-limiting step in protein folding, *Nat Struct Biol 6,* 554-556.

Polverino de Laureto, 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.

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.

Privalov, P. L. (1990) Cold denaturation of proteins, *Crit Rev Biochem Mol Biol 25,* 281-305.

Privalov, P. L., Griko Yu, V., Venyaminov, S., and Kutyshenko, V. P. (1986) Cold denaturation of myoglobin, *J Mol Biol 190,* 487-498.

Pucciarelli, S., Parker, S. K., Detrich, H. W., 3rd, and Melki, R. (2006) Characterization of the cytoplasmic chaperonin containing TCP-1 from the Antarctic fish *Notothenia coriiceps*, *Extremophiles 10,* 537-549.

Redl, B., Walleczek, J., Stoffler-Meilicke, M., and Stoffler, G. (1989) Immunoblotting analysis of protein-protein crosslinks within the 50S ribosomal subunit of *Escherichia coli*. A study using dimethylsuberimidate as crosslinking reagent, *Eur J Biochem 181,* 351-356.

Ropson, I. J., and Frieden, C. (1992) Dynamic NMR spectral analysis and protein folding: Identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by 19F NMR, *Proc Natl Acad Sci USA 89,* 7222-7226.

Saab-Rincon, G., Gualfetti, P. J., and Matthews, C. R. (1996) Mutagenic and thermodynamic analyses of residual structure in the α subunit of tryptophan synthase, *Biochemistry 35,* 1988-1994.

Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins, *J Am Chem Soc 132,* 8407-8418.

Sanchez, I. C. (1979) Phase-transition behavior of the isolated polymer-chain, *Macromol 12,* 980-988.

Sato, S., and Raleigh, D. P. (2002) pH-dependent stability and folding kinetics of a protein with an unusual α-β topology: The C-terminal domain of the ribosomal protein L9, *J Mol Biol 318,* 571-582.

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.

Sato, S., Luisi, D. L., and Raleigh, D. P. (2000) pH jump studies of the folding of the multidomain ribosomal protein L9: The structural organization of the N-terminal domain does not affect the anomalously slow folding of the C-terminal domain, *Biochemistry 39,* 4955-4962.

Schlorb, C., Ackermann, K., Richter, C., Wirmer, J., and Schwalbe, H. (2005) Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states, *J Biomol NMR 33,* 95-104.

Schlorb, C., Mensch, S., Richter, C., and Schwalbe, H. (2006) Photo-CIDNP reveals differences in compaction of non-native states of lysozyme, *J Am Chem Soc 128,* 1802-1803.

Scholtz, J. M., Grimsley, G. R., and Pace, C. N. (2009) Solvent denaturation of proteins and interpretations of the m value, *Methods Enzymol 466,* 549-565.

Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 A resolution, *Science 310,* 827-834.

Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426,* 900-904.

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.

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.

Shan, B., McClendon, S., Rospigliosi, C., Eliezer, D., and Raleigh, D. P. (2010) The cold denatured state of the C-terminal domain of protein L9 is compact and contains both native and non-native structure, *J Am Chem Soc 132,* 4669-4677.

Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability, *FASEB J 10,* 27-34.

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.

Stumpe, M. C., and Grubmuller, H. (2007) Interaction of urea with amino acids: Implications for urea-induced protein denaturation, *J Am Chem Soc 129*, 16126-16131.

Stumpe, M. C., and Grubmuller, H. (2008) Polar or apolar-The role of polarity for urea-induced protein denaturation, *Plos Comput Biol 4*.

Sung, Y. H., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family, *J Mol Biol 372*, 689-707.

Swint-Kruse, L., and Robertson, A. D. (1995) Hydrogen bonds and the pH dependence of ovomucoid third domain stability, *Biochemistry 34*, 4724-4732.

Tamura, A., Kimura, K., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .2. NMR Studies, *Biochemistry 30*, 11313-11320.

Tamura, A., Kimura, K., Takahara, H., and Akasaka, K. (1991) Cold denaturation and heat denaturation of streptomyces subtilisin inhibitor .1. CD and DSC Studies, *Biochemistry 30*, 11307-11313.

Tanford, C. (1968) Protein denaturation, *Adv Protein Chem 23*, 121-282.

Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Adv Protein Chem 24*, 1-95.

Tanford, C., Kawahara, K., and Lapanje, S. (1966) Proteins in 6 M guanidine hydrochloride. Demonstration of random coil behavior, *J Biol Chem 241*, 1921-1923.

Tang, Y., Goger, M. J., and Raleigh, D. P. (2006) NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain, *Biochemistry 45*, 6940-6946.

Tantos, A., Friedrich, P., and Tompa, P. (2009) Cold stability of intrinsically disordered proteins, *FEBS Lett 583*, 465-469.

Teilum, K., Kragelund, B. B., and Poulsen, F. M. (2002) Transient structure formation in unfolded acyl-coenzyme A-binding protein observed by site-directed spin labelling, *J Mol Biol 324*, 349-357.

Thomsen, J. K., Kragelund, B. B., Teilum, K., Knudsen, J., and Poulsen, F. M. (2002) Transient intermediary states with high and low folding probabilities in the apparent two-state folding equilibrium of ACBP at low pH, *J Mol Biol 318*, 805-814.

Tian, J., and Garcia, A. E. (2011) Simulations of the confinement of ubiquitin in self-assembled reverse micelles, *J Chem Phys 134*, 225101.

Todgham, A. E., Hoaglund, E. A., and Hofmann, G. E. (2007) Is cold the new hot? Elevated ubiquitin-conjugated protein levels in tissues of Antarctic fish as evidence for cold-denaturation of proteins *in vivo*, *J Comp Phys B 177,* 857-866.

Tollinger M, Crowhurst, K. A., Kay, L. E., and Forman-Kay, J. D. (2003) Site-specific contributions to the pH dependence of protein stability, *Proc Natl Acad Sci USA 100*, 4545-4550.

Tollinger, M., Skrynnikov, N. R., Mulder, F. 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.

Tran, H. T., and Pappu, R. V. (2006) Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions, *Biophys J 91*, 1868-1886.

Tran, H. T., Wang, X. L., and Pappu, R. V. (2005) Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins, *Biochemistry 44,* 11369-11380.

Tsai, C. J., Maizel, J. V., Jr., and Nussinov, R. (2002) The hydrophobic effect: A new insight from cold denaturation and a two-state water structure, *Crit Rev Biochem Mol Biol 37*, 55-69.

Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics, *Protein Sci 11,* 739-756.

Uversky, V. N. (2009) Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding, *Protein J 28,* 305-325.

Uversky, V. N. (2013) Unusual biophysics of intrinsically disordered proteins, *Biochim Biophys Acta 1834,* 932-951.

Uversky, V. N., Gillespie, J. R., Millett, I. S., Khodyakova, A. V., Vasiliev, A. M., Chernovskaya, T. V., Vasilenko, R. N., Kozlovskaya, G. D., Dolgikh, D. A., Fink, A. L., Doniach, S., and Abramov, V. M. (1999) Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic pH, *Biochemistry 38,* 15009-15016.

Uversky, V. N., Li, J., and Fink, A. L. (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation, *J Biol Chem 276,* 10737-10744.

Uzawa, T., Kimura, T., Ishimori, K., Morishima, I., Matsui, T., Ikeda-Saito, M., Takahashi, S., Akiyama, S., and Fujisawa, T. (2006) Time-resolved small-angle X-ray scattering investigation of the folding dynamics of heme oxygenase: Implication of the scaling relationship for the submillisecond intermediates of protein folding, *J Mol Biol 357,* 997-1008.

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.

Vendruscolo, M., and Paci, E. (2003) Protein folding: Bringing theory and experiment closer together, *Curr Opin Struc Biol 13,* 82-87.

Vitalis, A., and Pappu, R. V. (2009) ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions, *J Comput Chem 30,* 673-699.

Voelz, V. A., Singh, V. R., Wedemeyer, W. J., Lapidus, L. J., and Pande, V. S. (2010) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment, *J Am Chem Soc 132,* 4702-4709.

Voets, I. K., Cruz, W. A., Moitzi, C., Lindner, P., Areas, E. P. G., and Schurtenberger, P. (2010) DMSO-induced denaturation of hen egg white lysozyme, *J Phys Chem B 114,* 11875-11883.

Volkov, V., and Hamm, P. (2004) A two-dimensional infrared study of localization, structure, and dynamics of a dipeptide in membrane environment, *Biophys J 87,* 4213-4225.

Wagner, G. (1997) An account of NMR in structural biology, *Nature Struct Biol 4 Suppl,* 841-844.

Walleczek, J., Martin, T., Redl, B., Stoffler-Meilicke, M., and Stoffler, G. (1989) Comparative cross-linking study on the 50S ribosomal subunit from *Escherichia coli*, *Biochemistry 28,* 4099-4105.

Wang, Z., Plaxco, K. W., and Makarov, D. E. (2007) Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins, *Biopolymers 86,* 321-328.

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.

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.

Wirmer, J., Peti, W., and Schwalbe, H. (2006) Motional properties of unfolded ubiquitin: A model for a random coil protein, *J Biomol NMR 35,* 175-186.

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, *Angewan Chem 43,* 5780-5785.

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 222*, 311-333.

Wong, K. B., Freund, S. M., 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.

Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein, *Proc Natl Acad Sci USA 105*, 13367-13372.

Xue, Y., Podkorytov, I. S., Rao, D. K., Benjamin, N., Sun, H. L., and Skrynnikov, N. R. (2009) Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain, *Protein Sci 18*, 1401-1424.

Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Evdokimova, E., Liu, C. S., Semesi, A., Guido, V., Edwards, A. M., and Arrowsmith, C. H. (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins, *J Am Chem Soc 127*, 16512-16517.

Yi, Q., Scalley-Kim, M. L., Alm, E. J., and Baker, D. (2000) NMR characterization of residual structure in the denatured state of protein L, *J Mol Biol 299*, 1341-1351.

Yoo, T. Y., Meisburger, S. P., Hinshaw, J., Pollack, L., Haran, G., Sosnick, T. R., and Plaxco, K. (2012) Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state, *J Mol Biol 418*, 226-236.

Yoshidome, T., and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein, *Phys Rev E 79*, 030905.

Zhang, O. W., and FormanKay, J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions, *Biochemistry 36*, 3959-3970.

Zhang, O., and Formankay, J. D. (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, *Biochemistry 34*, 6784-6794.

Zhou, H. X. (2004) Polymer models of protein stability, folding, and interactions, *Biochemistry 43*, 2141-2154.

Zhou, H. X., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences, *Annu Rev Biophysics 37*, 375-397.

Zwanzig, R., Szabo, A., and Bagchi, B. (1992) Levinthal's paradox, *Proc Natl Acad Sci USA 89*, 20-22.

Appendix 1. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-^{1}H cross peaks in the HSQC spectra of the single Cys mutants of CTL9 at pH 2.0.

Residue	E61C	K74C	K96C	K109C	D119C	K149C
A59	0	0.72	0.89	1.04	1.18	1.11
E60	0	0.53	1.03	1.01	1.03	1.13
E61		0.44	0.90	0.81	0.92	1.14
L62			0.56	0.75	0.88	
A63		0.32	0.46	0.21	0.80	
N64	0	0.65	0.65	0.64	0.65	1.00
A65			0.56		0.92	0.80
K66	0	0.48	0.78	1.00	0.92	1.09
K67	0		0.88		0.92	1.11
L68	0		0.93		0.94	1.06
K69	0	0.27			0.65	1.03
E70	0.27	0.20		0.81	0.92	1.07
Q71	0.28	0	0.83	0.82	0.86	1.04
L72		0	1.00	0.69	0.83	
E73	0.79	0			0.77	0.94
K74			0.83		0.75	1.15
L75				0.57		
T76	0.74	0	0.77	0.60	0.81	1.06
V77					0.64	
T78	0.71	0	0.60	0.44	0.59	1.11
I79	0.76	0.21	0.48	0.50	0.64	1.11
P80						
A81	0.91	0.23	0.28	0.33	0.53	1.08
K82		0.20		0.61	0.66	1.10
A83	0.77	0.37	0.40	0.76	0.72	1.09
G84	0.97		0.22	0.53	0.62	1.17
E85	0.79	0.43	0.25	0.72	0.76	1.04
G86	0.86	0.38	0.26	0.59	0.68	1.10
G87	0.96		0.19		0.84	1.09
R88		0.35	0	0.60	0.60	1.10
L89		0.28	0	0.48	0.34	
F90		0.23		0.35	0.41	1.02
G91	0.75	0.17	0.10	0.29	0.30	1.11
S92	0.85	0.24	0.14	0.47	0.53	1.10
I93		0.30	0	0.35	0.30	
T94	0.83	0.31	0.16	0.47	0.40	
S95	0.76	0.29	0	0.37	0.34	1.00
K96		0.42			0.55	
Q97		0.21		0.29	0.43	1.04

Residue	E61C	K74C	K96C	K109C	D119C	K149C
I98	0.78		0		0.4	1.10
A99	0.70		0	0.26	0.11	1.04
E100	0.78		0		0	1.04
S101		0.26	0	0.35	0	1.05
L102		0.64	0		0.31	
Q103	0.78	0.43	0.30		0.36	1.12
A104	0.46	0.40	0.57	0.30	0.37	0.80
Q105	0.90	0.29	0.29	0.28	0.32	1.09
H106	0.66	0.62	0.44	0	0.64	1.10
G107	0.91	0.57	0.37	0	0.48	1.14
L108	0.82	0.36	0.42	0	0.33	1.08
K109		0.39			0.42	0.81
L110						
D111	0.79	0.27	0.24	0	0.23	0.99
K112		0.40		0	0.41	
R113		0.48	0.57	0	0.2	
K114						1.09
I115	0.76		0.44	0.71	0.21	0
E116		0.37	0.54	0.35	0	1.06
L117			0.90			
A118	0.91	0.44	0.80	0.73	0	1.05
D119	0.90	0.38	0.90	0.33		1.07
A120	0.90	0.32	0.62	0.30		1.06
I121	0.94	0.30	0.58	0.22	0	1.04
R122	1.10	0.39	0.64	0.74	0	1.02
A123	1.10		0.90	0.74	0	1.02
L124	0.82		0.61	0.21		
G125	0.79	0.27	0.62	0.82	0	0.96
Y126	0.91	0.41	0.75	0.44	0.24	1.05
T127	0.73	0.38	0.81	0.45	0.19	1.01
N128	0.78	0.45	0.95	0.45	0.28	
V129	0.88	0.57	0.76	0.74	0.58	1.02
P130						
V131				0.486	0.39	
K132	0.88	0.44	0.78	0.56	0.36	
L133		0.40		0.68	0.43	
H134	0.79	0.52	0.92	0.63	0.36	0.76
P135						
E136	0.80	0.44	0.96	0.54	0.32	0.49
V137		0.80	0.47	0.84	0.85	1.09
T138	0.87	0.82	0.76	0.82	0.92	0.73
A139	0.86	0.71	0.80	0.82	0.78	0.56

Residue	E61C	K74C	K96C	K109C	D119C	K149C
T140	0.83	0.74	0.79	0.94	0.91	0.69
L141	0.81	0.76	0.76	0.74		0.98
K142		0.82	0.82	0.91	1.01	0
V143	0.72	0.88	0.94	0.79	0.94	0.59
H144	0.82	0.78	0.94	0.96	0.89	0.35
V145			0.92			0
T146	0.89	0.92	0.98	0.91	0.94	0.31
E147	0.74	0.93	0.95	1.00	1.04	0
Q148	0.86	0.84	1.06	0.99	1.01	0
K149	0.79	0.85	0.94	1.15	0.96	

Appendix 2. Intensity ratio (I_{para}/I_{dia}) of the ^{15}N-^{1}H cross peaks in the HSQC spectra of the single Cys mutants of CTL9 at pH 2.0.

Residue	E61C	K74C	K96C	K109C	D119C	K149C
A59	0	1.23	1.08	0.99	1.05	1.02
E60	0	0.84	1.06	1.03	0.99	1.13
E61		0.82	1.05	1.13	1.01	1.10
L62						
A63					1.03	
N64	0	0.77	0.97	0.94	1.02	1.06
A65	0	0.94	0.92	1.04	1.00	1.06
K66	0	0.78	1.01	0.98	1.19	
K67	0		1.02	1.09	1.09	1.11
L68						1.04
K69	0	0.42	0.96	0.97	0.98	0.97
E70	0.33	0.26	0.93	0.99	1.08	0.99
Q71	0	0	0.77	0.97	1.118	0.94
L72						
E73	0.78	0		0.36	1.09	1.04
K74			0.90			0.98
L75						
T76	0.88		0.76	0.98	1.12	1.10
V77	0.99	0	0.67	1.10	1.16	1.16
T78	0.82	0	0.74	0.85	1.14	0.96
I79	0.83	0	0.55	0.90	1.06	
P80						
A81	0.75	0			0.99	
K82	0.95	0.89	0.37	0.94	0.98	1.08
A83						
G84	0.94	0.75	0. 30	1.04	1.04	1.07
E85	0.88	0.89	0.34	1.09	1.11	1.05
G86	1.01	0.83	0.33	1.05	0.96	1.10
G87	0.81	0.74	0.24	0.94	1.05	1.04
R88	0.80	0.92	0.29	0.96	1.12	1.08
L89	0.85		0.67	0.84	0.79	
F90	0.94	0.76	0	0.96	0.98	1.00
G91	0.77	0.58	0.16	0.87	0.94	1.05
S92	0.91	0.64	0.17	0.84	1.06	1.08
I93	0.81	0.76	0.20	0.96	1.06	1.09
T94	0.87	0.79	0	0.97	0.95	0.98
S95	0.84	0.83	0	0.83		1.04
K96	0.85	0.85		0.83	1.04	1.07
Q97	0.87	0.85	0	0.78	1.09	1.19

Residue	E61C	K74C	K96C	K109C	D119C	K149C
I98	0.90	0.48	0		1.04	1.04
A99	0.85	0.77	0	0.50	0.92	1.01
E100	0.93	0.79	0	0.43	0.73	1.03
S101	0.94		0	0.48	0.93	1.09
L102						
Q103						
A104	0.76	0.91	0.76	0.90		0.84
Q105	0.96		0.32	0.34	0.79	1.15
H106	0.96	0.89	0.57	0.26	1.06	1.14
G107	0.84	0.88	0.55	0	0.83	1.10
L108	0.82	1.03	0.49		0.84	1.05
K109	0.99	1.05	0.68		0.95	1.04
L110					1.12	
D111	0.92	0.87	0.42		0.74	1.16
K112	1.02	0.91	0.65	0		
R113			0.65		0.86	
K114	0.97	0.89	1.08	0	0.53	0.98
I115	0.85	0.80	0.67	0		
E116	0.97	0.80	0.95	0.96	0	1.16
L117						
A118					0	
D119	0.82	0.95	0.81	0.94		1.16
A120						
I121	0.77	0.92	0.86	0.75	0	1.01
R122	0.96	0.91	0.75	0.37	0	1.05
A123	0.90			0.88		
L124	0.85	0.98	0.77	0.72	0	1.07
G125	0.85	1.05	0.82	0.78	0.32	1.06
Y126	0.82	1.01	0.99	0.95	0.40	1.08
T127	0.81	0.96	1.03	0.97	0.54	1.10
N128	0.70	0.96	0.79	0. 90	0.72	1.02
V129	0.89	0.98	1.02	0.85	0.63	1.11
P130						
V131	0.77	0.89	0.95	0.96	0.51	1.01
K132	0.97	0.90	0.94	1.00		0.92
L133						
H134	0.84	0.98	0.97	1.00	0.67	1.02
P135						
E136	0.76	1.05	0.99	0.53	0.58	0.61
V137	1.01	1.05	0.74	0.93	0.94	0.86
T138	0.92	1.01	0.98	1.03	0.98	0.93
A139	0.77	0.92	0.99	1.04	0.90	0.68

Residue	E61C	K74C	K96C	K109C	D119C	K149C
T140	0.80	1.03	1.03	1.10	1.06	0.77
L141	0.86	0.83	0.97	0.37	0.94	0.98
K142	0.86	1.00				0.84
V143	0.94	1.03	0.96	0.92	1.08	0.70
H144	0.96	0.93	0.98	1.02	0.96	0.40
V145	0.89	0.90	1.01	1.07	1.01	0.55
T146	0.92	0.92	0.93	1.08	1.04	0.33
E147		1.08	0.97	0.91	0.96	0
Q148	0.89	1.02	1.00	0.93	0.95	0
K149	0.93	1.03	1.08	1.08	1.06	

Appendix 3. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the E61C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.60	123.68	8.40	123.74
E60	8.48	120.51	8.24	120.10
E61C				
L62				
A63				
N64	8.23	117.36	8.09	117.52
A65			7.92	124.07
K66	8.15	119.53	8.04	120.30
K67	8.10	121.48	8.07	122.58
L68	8.09	122.57		
K69	8.18	121.48	8.25	122.61
E70	8.18	120.50	8.17	121.72
Q71	8.32	121.62	8.31	122.16
L72				
E73	8.17	120.99	8.08	120.38
K74	8.20	122.25		
L75				
T76	8.18	116.34	8.09	116.30
V77			8.00	122.57
T78	8.30	119.78	8.09	118.78
I79	8.30	126.05	8.08	124.95
P80				
A81	8.36	124.38	8.12	124.42
K82	8.22	120.65	8.04	120.58
A83	8.30	125.47		
G84	8.38	108.33	8.14	108.13
E85	8.24	119.76	8.03	119.42
G86	8.53	110.02	8.27	109.72
G87	8.26	108.62	8.02	108.30
R88			7.97	120.40
L89			8.07	123.22
F90	8.22	121.02	8.13	120.98
G91	8.30	110.36	8.12	110.12
S92	8.16	115.71	8.00	115.51
I93	8.25	122.59	8.09	122.38
T94	8.19	117.59	7.99	117.28

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.31	117.50	8.08	118.00
K96			8.14	123.23
Q97			8.13	121.40
I98	8.13	122.37	7.99	122.43
A99	8.27	126.74	8.13	127.56
E100	8.27	119.27	8.09	119.72
S101			8.12	116.68
L102				
Q103	8.15	119.97		
A104	8.11	124.06	8.08	125.56
Q105	8.16	118.53	8.05	119.09
H106	8.39	118.61	8.31	118.94
G107	8.38	109.47	8.22	109.65
L108	8.06	121.82	7.93	121.75
K109	8.33	122.48	8.29	123.20
L110				
D111	8.42	120.38	8.38	120.76
K112	8.25	122.74	8.15	122.48
R113	8.20	121.74		
K114			8.25	123.67
I115	8.11	122.81	8.07	123.22
E116			8.12	125.34
L117				
A118	8.32	124.52		
D119	8.25	116.69	8.15	117.84
A120	8.08	124.36		
I121	7.96	119.57	7.82	119.89
R122	8.19	125.13	8.24	125.03
A123			8.15	124.70
L124	8.13	121.37	8.00	121.68
G125	8.25	108.82	8.08	108.87
Y126	7.96	119.87	7.84	119.76
T127	8.04	116.03	7.92	115.50
N128	8.36	121.63	8.20	121.42
V129	8.05	122.27	7.84	121.24
P130				
V131			7.94	120.75
K132	8.33	125.46	8.15	125.35
L133				

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.50	119.37	8.46	119.53
P135				
E136	8.53	121.26	8.35	121.30
V137	8.28	122.40	8.09	121.83
T138	8.25	118.90	8.05	118.15
A139	8.36	127.10	8.10	126.54
T140	8.12	114.16	7.92	113.75
L141	8.21	124.79	7.99	124.58
K142	8.31	123.02	8.20	123.10
V143	8.00	121.01	7.91	121.05
H144	8.65	122.88	8.50	122.82
V145	8.28	123.36	8.13	122.94
T146	8.33	119.15	8.14	118.66
E147	8.41	123.42		
Q148	8.45	122.54	8.28	122.32
K149	8.49	123.92	8.28	123.54

Appendix 4. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the K74C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.59	124.54	8.33	122.19
E60	8.44	120.49	8.14	119.45
E61	8.40	122.26	8.19	120.89
L62	8.25	124.44		
A63			7.87	122.43
N64	8.27	117.28	8.06	117.72
A65	8.09	123.35		
K66	8.10	119.73		
K67	8.01	121.71	8.02	121.17
L68	8.01	122.59	7.97	119.74
K69	8.11	121.44	8.14	121.67
E70	8.12	120.33	8.11	120.71
Q71	8.27	121.22	8.21	121.12
L72	8.12	123.45		
E73	8.16	120.48	8.01	119.70
K74C				
L75				
T76	8.14	115.42		
V77				
T78	8.24	119.40	7.99	118.39
I79	8.24	126.64	7.99	122.96
P80				
A81	8.30	125.13		
K82	8.20	120.34	7.94	120.66
A83	8.26	126.09		
G84	8.35	107.07	8.06	110.78
E85	8.23	123.97	7.95	118.99
G86	8.51	108.93	8.19	111.92
G87	8.23	107.46	7.95	110.88
R88	8.06	120.69	7.90	119.73
L89	8.19	124.31	8.01	121.80
F90	8.18	121.02	8.07	120.15
G91	8.28	109.27	8.05	112.23
S92	8.07	115.00	7.92	116.18
I93			8.08	121.25
T94	8.14	117.03	7.91	117.41

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.30	117.73	8.00	117.96
K96	8.21	123.96	8.06	121.80
Q97	8.14	120.80	8.06	120.44
I98	8.07	122.44	7.91	121.20
A99	8.21	126.90	8.05	124.96
E100	8.29	119.15	8.02	119.24
S101	8.17	116.28	8.05	117.01
L102				
Q103	8.07	119.76		
A104	8.05	124.37	8.01	123.50
Q105	8.10	118.31	7.98	118.80
H106	8.34	118.39	8.24	118.66
G107	8.36	108.31	8.15	111.89
L108			7.86	120.70
K109	8.29	122.60	8.22	121.88
L110				
D111			8.32	119.95
K112	8.20	122.36	8.08	121.25
R113	8.14	121.44		
K114	8.14	122.21	8.18	123.10
I115	8.05	122.74	8.01	121.39
E116	8.27	123.54	8.08	123.37
L117	8.27	125.09		
A118				
D119	8.29	116.94	8.06	117.72
A120				
I121	7.92	119.16	7.77	119.26
R122	8.10	124.67	8.14	123.10
A123	8.10	124.98		
L124	8.02	120.91		
G125	8.17	107.26	8.01	111.33
Y126	7.89	119.54	7.77	119.26
T127	8.00	115.30	7.85	116.13
N128	8.34	121.80	8.13	120.42
V129	8.01	122.23	7.76	120.27
P130				
V131			7.88	120.01
K132	8.30	126.16	8.08	123.37
L133				

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.44	119.20	8.40	119.08
P135				
E136	8.50	121.48	8.28	120.38
V137	8.24	122.68	8.02	120.76
T138	8.21	118.70	7.97	118.04
A139	8.33	128.07	8.02	124.17
T140	8.08	113.43	7.84	114.87
L141	8.15	125.90	7.92	122.78
K142				
V143	7.98	121.43	7.86	120.27
H144	8.61	123.45	8.42	121.52
V145			8.06	121.60
T146	8.36	119.88	8.06	118.45
E147	8.38	124.07		
Q148	8.42	123.00	8.20	121.37
K149	8.47	124.64	8.19	122.14

Appendix 5. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the K96C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.61	123.72	8.40	123.76
E60	8.46	120.28	8.21	120.02
E61	8.44	121.94	8.25	122.00
L62	8.29	124.06		
A63				
N64	8.29	117.37	8.13	117.64
A65				
K66	8.14	119.47	8.03	120.29
K67	8.05	121.44	8.07	122.57
L68	8.09	122.52		
K69	8.18	121.45	8.28	123.19
E70	8.19	120.48	8.17	121.74
Q71	8.32	121.24	8.31	122.17
L72				
E73	8.17	121.00	8.07	120.30
K74	8.20	122.23	8.20	123.09
L75				
T76	8.18	116.34	8.09	116.30
V77			8.00	122.53
T78	8.31	119.77	8.08	118.74
I79	8.36	126.02	8.07	124.91
P80				
A81	8.30	124.35		
K82	8.22	120.61	8.04	120.56
A83	8.30	125.48		
G84	8.38	108.33	8.13	108.11
E85	8.24	119.76	8.02	119.41
G86	8.53	110.02	8.26	109.70
G87	8.26	108.59	8.02	108.27
R88	8.09	120.40	7.96	120.39
L89			8.07	123.27
F90	8.24	121.06	8.17	120.80
G91	8.29	110.36	8.11	110.10
S92	8.16	115.66	7.99	115.48
I93			8.09	122.32
T94	8.20	117.55	7.98	117.09

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.31	117.97	8.08	117.70
K96C				
Q97			8.13	120.95
I98	8.11	122.16	7.93	122.07
A99	8.29	126.99	8.13	127.65
E100	8.27	119.35	8.08	119.72
S101	8.25	116.50	8.11	116.56
L102				
Q103	8.15	119.89		
A104	8.11	123.99	8.08	125.56
Q105	8.16	118.46	8.04	119.05
H106	8.39	118.51	8.30	118.90
G107	8.38	109.43	8.22	109.61
L108	8.05	121.80	7.92	121.73
K109	8.34	122.46		
L110				
D111	8.42	120.53	8.38	120.77
K112			8.15	122.46
R113	8.21	121.64	8.15	122.46
K114			8.26	123.57
I115	8.10	122.75	8.07	123.27
E116	8.36	124.35	8.12	125.32
L117				
A118	8.36	124.34		
D119			8.15	117.85
A120	8.08	124.35		
I121	7.96	119.55	7.82	119.90
R122	8.19	125.11	8.24	125.03
A123	8.21	125.28		
L124	8.14	121.36	8.00	121.69
G125	8.25	108.80	8.08	108.85
Y126	7.96	119.85	7.84	119.76
T127	8.04	116.04	7.92	115.49
N128	8.36	121.63	8.21	121.43
V129	8.06	122.29	7.83	121.21
P130				
V131			7.94	120.74
K132	8.34	125.47	8.14	125.35
L133				

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.51	119.36	8.47	119.54
P135				
E136	8.53	121.26	8.35	121.29
V137	8.28	122.43	8.08	121.82
T138	8.25	118.91	8.04	118.13
A139	8.31	127.10	8.10	126.52
T140	8.13	114.17	7.91	113.74
L141	8.21	124.77	7.98	124.51
K142	8.32	123.02		
V143	8.00	121.01	7.91	121.04
H144	8.66	122.87	8.49	122.81
V145	8.28	123.80	8.12	122.94
T146	8.33	119.13	8.13	118.64
E147	8.41	123.41	8.21	123.22
Q148			8.27	122.31
K149	8.47	124.15	8.26	123.57

Appendix 6. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the K109C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.59	124.55	8.34	123.73
E60	8.44	120.50	8.15	119.99
E61	8.40	122.24	8.19	121.96
L62				
A63				
N64	8.27	117.30	8.07	117.63
A65	8.08	123.28	7.88	124.08
K66	8.10	119.64	7.98	120.32
K67	7.98	121.46	8.01	123.21
L68	8.00	122.50		
K69	8.10	121.33		
E70	8.12	120.05	8.11	121.71
Q71	8.24	120.98	8.22	122.26
L72	8.19	123.39		
E73	8.14	120.34	8.01	120.30
K74	8.13	121.87	8.07	123.19
L75				
T76	8.09	115.24	8.02	116.21
V77			7.92	122.37
T78	8.30	119.14	8.01	118.67
I79	8.25	126.73	8.01	124.85
P80				
A81	8.30	125.22	8.03	123.96
K82	8.17	120.56	7.98	120.32
A83	8.26	126.14		
G84	8.35	107.08	8.07	108.10
E85	8.20	119.93	7.96	119.35
G86	8.50	108.95	8.20	109.66
G87	8.23	107.47	7.95	108.24
R88	8.06	120.70	7.91	120.36
L89	8.20	123.73		
F90	8.18	121.02	8.07	120.93
G91	8.28	109.26	8.05	110.08
S92	8.14	115.42	7.93	115.44
I93				
T94			7.92	117.18

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.30	117.72	8.01	117.93
K96				
Q97	8.14	120.78	8.06	121.33
I98			8.02	122.62
A99	8.21	126.89	8.06	127.52
E100	8.25	119.47	8.02	119.66
S101	8.17	116.25	8.05	116.62
L102				
Q103	8.07	119.75		
A104	8.05	124.36	8.01	125.51
Q105	8.10	118.35	7.99	119.06
H106	8.34	118.40	8.24	118.90
G107	8.37	108.27	8.16	109.62
L108			7.91	121.59
K109C				
L110				
D111	8.38	120.60	8.29	121.23
K112			8.03	122.33
R113	8.19	124.30	8.06	122.26
K114	8.21	123.01	8.17	123.53
I115	8.04	122.55		
E116			8.08	125.34
L117	8.25	124.41		
A118	8.27	125.05		
D119	8.29	116.88	8.07	117.62
A120	8.00	124.90		
I121	7.92	119.07	7.78	119.72
R122	8.13	123.88	8.15	125.23
A123	8.11	124.68	8.08	124.54
L124	8.07	121.87	7.94	121.62
G125	8.16	107.15	8.01	108.84
Y126	7.89	119.95	7.78	119.72
T127	8.00	115.26	7.86	115.42
N128	8.33	121.79	8.15	121.40
V129	7.96	122.38	7.77	121.12
P130				
V131			7.88	120.74
K132	8.30	126.15	8.08	125.34
L133				

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.44	119.20	8.41	119.46
P135				
E136	8.50	121.48	8.25	122.11
V137	8.25	122.68	8.02	121.77
T138	8.21	118.70	7.97	118.05
A139	8.33	128.06	8.03	126.46
T140	8.08	113.43	7.85	113.69
L141	8.15	125.90	7.93	124.55
K142	8.26	123.55	8.14	123.07
V143	8.00	120.79	7.86	121.10
H144	8.62	123.45	8.43	122.81
V145	8.23	123.95		
T146	8.34	119.01	8.07	118.61
E147	8.38	124.07	8.07	122.94
Q148	8.42	123.00	8.20	122.63
K149	8.47	124.60	8.22	123.35

Appendix 7. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the D119C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.58	123.76	8.50	121.14
E60	8.44	120.25	8.33	124.22
E61	8.40	121.82	8.34	118.40
L62				
A63			8.09	114.03
N64	8.26	117.37	8.25	122.24
A65	8.10	123.94		
K66	8.11	118.36	8.14	120.97
K67	8.00	121.28	8.18	120.81
L68	8.08	122.03		
K69			8.35	108.33
E70	8.13	120.02	8.30	122.22
Q71			8.46	119.15
L72	8.14	122.35		
E73	8.14	120.26	8.14	122.35
K74	8.11	121.71		
L75				
T76	8.11	115.83	8.18	120.39
V77			8.11	122.95
T78	8.25	120.81	8.18	116.52
I79	8.26	125.87	8.19	123.64
P80				
A81	8.33	124.22		
K82	8.18	120.39		
A83	8.26	125.31		
G84	8.35	108.33	8.25	123.74
E85	8.23	124.29	8.13	120.02
G86	8.50	109.99	8.38	123.40
G87	8.22	108.63	8.13	115.70
R88	8.06	120.38	8.10	119.49
L89				
F90	8.24	123.30	8.26	119.50
G91	8.27	110.31	8.22	126.21
S92	8.13	115.70	8.10	121.13
I93			8.20	119.71
T94	8.14	117.29	8.10	123.94

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.28	117.86	8.15	121.51
K96			8.26	124.28
Q97	8.14	120.97	8.23	124.29
I98	8.06	122.46		
A99	8.22	126.21	8.25	120.81
E100	8.26	119.50	8.19	122.85
S101	8.18	116.52	8.22	108.63
L102				
Q103	8.10	119.48		
A104	8.06	123.78	8.21	118.73
Q105	8.16	117.68	8.14	120.26
H106	8.34	118.4	8.44	120.25
G107	8.35	109.41	8.33	126.99
L108	8.01	121.74	8.06	123.78
K109	8.29	122.22	8.42	122.48
L110				
D111	8.37	119.90	8.50	109.99
K112	8.22	122.09		
R113	8.19	123.64	8.27	125.31
K114			8.35	109.41
I115	8.05	122.19	8.16	117.68
E116	8.30	124.42	8.24	123.30
L117	8.25	123.74		
A118	8.26	124.28		
D119C				
A120				
I121	7.97	119.74	8.01	122.12
R122	8.19	124.72	8.37	119.90
A123	8.16	125.03	8.30	119.08
L124	8.10	121.13	8.11	118.36
G125	8.22	108.63	8.16	125.03
Y126	7.92	119.81	8.01	121.74
T127	8.00	115.87	8.06	120.38
N128	8.33	121.47	8.30	125.35
V129	8.01	122.12	8.00	121.28
P130				
V131			8.08	122.09
K132	8.30	125.35	8.30	124.42
L133			8.26	125.31

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.46	119.15	8.58	123.76
P135				
E136	8.50	121.14	8.47	123.89
V137	8.25	122.24	8.19	124.72
T138	8.21	118.73	8.14	117.29
A139	8.33	126.99	8.22	122.09
T140	8.09	114.03	8.05	122.19
L141			8.11	121.71
K142	8.26	122.95		
V143	7.98	121.06	8.06	122.46
H144	8.62	122.86	8.62	122.86
V145			8.26	122.95
T146	8.30	119.08	8.26	117.37
E147	8.38	123.40	8.42	122.48
Q148	8.42	122.48	8.40	121.82
K149	8.47	123.89		

Appendix 8. List of the chemical shifts obtained from the ^{15}N-HSQC spectra for the K149C-CTL9 DSE in acid and urea.

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
A59	8.61	123.72	8.40	123.74
E60	8.47	120.36	8.21	120.03
E61	8.44	121.99	8.26	122.01
L62				
A63				
N64	8.31	117.59	8.14	117.64
A65	8.11	124.04	7.94	124.08
K66	8.13	119.52		
K67	8.04	121.48	8.07	122.57
L68	8.09	122.52	8.04	120.28
K69	8.14	121.39	8.26	122.62
E70	8.20	120.73	8.18	121.75
Q71	8.32	121.26	8.31	122.18
L72				
E73	8.19	121.11	8.08	120.40
K74	8.16	121.45	8.16	122.47
L75				
T76	8.18	116.47		
V77			8.01	122.58
T78	8.31	119.97	8.09	118.79
I79	8.32	126.20		
P80				
A81	8.34	124.60		
K82	8.23	120.74	8.05	120.60
A83	8.31	125.57		
G84	8.38	108.38	8.14	108.11
E85	8.24	119.77	8.03	119.41
G86	8.53	110.06	8.27	109.69
G87	8.26	108.63	8.02	108.27
R88	8.09	120.42	7.97	120.39
L89				
F90	8.22	121.08	8.13	120.99
G91	8.30	110.41	8.12	110.09
S92	8.16	115.72	8.00	115.50
I93			8.10	122.38
T94			7.99	117.29

Residue	Acid induced DSE		Urea induced DSE	
	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
S95	8.31	118.11	8.08	118.00
K96			8.14	123.23
Q97	8.18	120.50	8.13	121.42
I98	8.14	122.49	7.99	122.46
A99	8.27	126.75	8.13	127.60
E100	8.28	119.33	8.09	119.76
S101	8.25	116.77	8.13	116.70
L102				
Q103	8.14	120.00		
A104	8.11	124.04	8.08	125.57
Q105	8.16	118.56	8.06	119.10
H106	8.39	118.63	8.31	118.96
G107	8.38	109.48	8.22	109.64
L108	8.05	121.82	7.93	121.74
K109	8.29	122.49	8.29	123.22
L110				
D111	8.42	120.53	8.39	120.87
K112				
R113	8.20	121.77		
K114	8.19	122.27	8.25	123.66
I115				
E116	8.36	124.41	8.12	125.34
L117				
A118	8.30	124.31		
D119	8.29	117.41	8.16	117.90
A120	8.07	124.40		
I121	7.95	119.66	7.82	119.91
R122	8.19	125.09	8.25	125.06
A123	8.19	125.20		
L124			8.01	121.70
G125	8.25	108.84	8.08	108.85
Y126	7.95	119.89		
T127	8.03	116.16	7.93	115.51
N128	8.45	121.28	8.21	121.46
V129	8.05	122.39	7.84	121.24
P130				
V131			7.94	120.74
K132			8.15	125.35
L133				

	Acid induced DSE		Urea induced DSE	
Residue	^{1}H (ppm)	^{15}N (ppm)	^{1}H (ppm)	^{15}N (ppm)
H134	8.50	119.36	8.47	119.53
P135				
E136	8.53	121.29	8.35	121.28
V137	8.26	122.65	8.09	121.82
T138	8.25	119.00	8.05	118.13
A139	8.36	127.14	8.10	126.50
T140	8.12	114.22	7.92	113.70
L141	8.20	124.87	8.00	124.59
K142			8.21	123.20
V143	8.00	121.11	7.91	121.02
H144	8.65	122.85	8.49	122.73
V145			8.13	122.90
T146	8.32	118.98	8.12	118.42
E147	8.39	123.28	8.19	122.85
Q148	8.49	122.36	8.30	122.03
K149C				