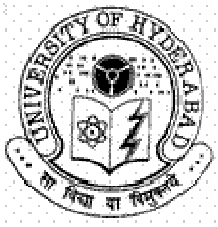
Stability, Chain Dynamics and Dimensions of End-States Across the Folding Equilibrium: Studies on Lysozyme, a Plant Phloem Protein *At*PP1 and Cytochrome *c* by NMR and Laser-Based Methods

A Thesis
Submitted for the Degree of

DOCTOR OF PHILOSOPHY

by

U MAHAMMAD YASIN



SCHOOL OF CHEMISTRY
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046
INDIA
December-2014

Dedicated to

My Family

CONTENTS

Chapter 3

Chapter 4

School of Chemistry
University of Hyderabad
Central University P. O.,
Hyderabad 500 046
India

STATEMENT

I hereby declare that the work embodied in this dissertation is the result of the investigation carried out by me in the School of Chemistry, University of Hyderabad, Hyderabad, under the supervision of **Prof. Abani K. Bhuyan.**

U MAHAMMAD YASIN

December 2014

STATEMENT VERIFIED

(Prof. ABANI K. BHUYAN)

PROJECT SUPERVISOR

School of Chemistry
University of Hyderabad
Central University P. O.,
Hyderabad 500 046
India

CERTIFICATE

Certified that the work embodied in this thesis entitled **'Stability, Chain Dynamics and Dimensions of End-States Across the Folding Equilibrium: Studies on Lysozyme, a Plant Phloem Protein *At*PP1 and Cytochrome *c* by NMR and Laser-Based Methods'** has been carried out by Mr. U Mahammad Yasin, under my supervision and the same has not been submitted elsewhere for a Degree.

Prof. ABANI K. BHUYAN
(THESIS SUPERVISOR)

Dean
School of Chemistry

Acknowledgements

I express my deep sense of gratitude and profound thanks to Prof. Abani K. Bhuyan for all of his support and guidance over the years. It has been an honour to work with and learn from him.

I would like to thank the present and former Dean(s), School of Chemistry, for their constant support, inspiration and for providing good facilities. I am extremely appreciative individually to all the faculty members of the school for their help, cooperation and encouragement at various stages.

I would like to express my sincere thanks to all of my colleagues: P Sashikantha Reddy, Dr. S Babu, D Ramakrishna, Hima Girija, Hariharan, G Mahender, Sony Reddy, Gaurang Chakravarthy for making the lab such a peaceful place. I would like to thank my seniors Dr. Prakash Prabhu, Dr. P Nageswar Rao for their valuable suggestions.

I thank all the non-teaching staff of the School of Chemistry for their assistance on various occasions. I specially thanks to Satyanarayana, Vijaya Lakshmi, Turabuddin, Durgesh, Venkateswarulu NMR facility.

I am thankful to NMR operators at TIFR Bombay, Prof. H S Atreya, IISc, Bangalore, Dr. Kiran S IICT, Hyderabad.

I am thankful to my School of Chemistry friends and colleagues Thiru, Pavan, Kishore, Shiva, Sudhir, Bhanu, Gangadhar, Chandu, Hari, Gupta, Durga Prasad, Ashok, Chandu, Rajgopal Reddy, Rajgopal Rao, Naga, Shiva Ranjan, Ramu Yadav, B Ramakrishna, Usha Rani, Ramesh, Pilli Kishore, Ganesh, GS Reddy, Nanda Kishore, Obaiah, Satish, Vikranth, Karunakar, Naveen Yadav,

Madhava Chari, V Venkat Rao, Suresh, A Suresh, Mahesh, Purushotham, TIFR friends Raji akka, Hema Chandra, Jitendar, Sharat, Ali, Vijaya Lakshmi, Susmitha, IISC friends Venki Reddy, Somnath Mandal, Lokeswar, Sachin, Shivanand.

I am also thankful to my MSc and Degree class mates Ramesh Niverthi, C Shiva Reddy, A Giri, C Siddha Reddy, Nagi Reddy, Shiva Shankar, Suresh Naik.

I cannot find words to express my feeling about my mother T K Shariefa Bee, father Dastagiri, my brothers Abdul Rasheed, Abdul Rahim, vodinas Noorjahan, Ameena, and their children Mehataj, Baba Dastagiri, and little girl Sajda and her brother, my atta, mama, pinnamma garu for their love and affection.

ABBREVIATIONS

1. GdnHCl, guanidinium hydrochloride
2. GdnDCl, guanidinium deutrochloride
3. CD, circular dichroism
4. Cyt *c*, cytochrome *c*
5. Ferricyt *c*, ferricytochrom *c*
6. Ferrocyt *c*, ferrocytochrome *c*
7. Carbonmonoxycyt *c*, the carbonmonoxide complex of ferrocytochrome *c*
8. *At*PP1, *Arabidopsis thaliana* phloem protein 1
9. GlcNAc, N-Acetyl-D-Glucosamine
10. sLED-PFG, stimulated echo longitudinal encode-decode pulse field gradient
11. FWHM, full width at half maximum
12. IDPs, intrinsically disorder proteins
13. SDS, sodium dodecyl sulfate
14. PAGE, poly acrylamide gel electrophoresis
15. $T_{1,}$ T_2, NOE, longitudinal relaxation, transverse relaxation, nuclear overhauser effect
16. CPMG, Carr-Purcell-Meiboom-Gill sequence
17. S_{conf}, conformational entropy
18. LTEC, Linear thermal expansion coefficients
19. ANS, 8-Anilinonaphthalene-1-sulfonic acid
20. $<R_H>$, average hydrodynamic radius

21. $K_{app,}$ apparent equilibrium constant
22. K_U, equilibrium constant for global unfolding
23. ΔG_D, Gibbs free energy of denaturation
24. $\Delta G°=\Delta G(H_2O)$, Gibbs free energy of denaturation in the absence of denaturant (unfolding free energy)
25. G_T , Thermal denaturation free energy
26. cP, centipoise
27. NCO, Native state of carbonmonoxycytochrome *c*
28. UCO, Unfolded state of carbonmonoxycytochrome *c*
29. LFEM, linear free energy extrapolation method
30. LFER, linear free energy relation
31. TOF, time of flight
32. Mb, Myoglobin
33. MbCO, Carboxymyoglobin
34. k_{on} , CO rebinding rate with myoglobin
35. k_{off} , CO replacement rate in carboxymyoglobin

$\Delta G°$ and $\Delta G(H_2O)$ refer same meaning and these two symbols are used interchangeably

***Symbols for different parameters are explained fully in the text**

1.1 Introduction

The protein folding problem is a difficult one, mainly because it is a many-body problem in aqueous solvent involving hundreds to thousands of atoms engaged in both protein intramolecular and protein-solvent intermolecular interactions by a variety of weak forces. The unfolded chain somehow collapses to acquire these interactions to give rise to secondary and tertiary structures whose amount and architectural details vary from one protein to another. The structured protein thrives with a range of less understood motional features at global, subglobal, and atomic levels that occur on different timescales ranging from picoseconds to seconds with varying frequency and amplitude. Whereas inherent motions provide functional abilities to proteins, the folded native state is remarkably compact - mean packing densities for all atom types fall between 0.7 and 0.8.[1] The packing density, which is the ratio of the volume enclosed by van der Waals envelope of an atom or group of atoms to the corresponding volume of space actually occupied, is thus closely comparable to those for small-molecule organic crystals[1] and metals like calcium, aluminium and lead. Consistent with such packing efficiencies are several reports of small intrinsic compressibility of globular proteins, which suggests a solid-like interior[2-4] but contrasts the large-scale subglobal fluctuations implying a liquid-like interior. As a synthesis of solid- and liquid-like behaviors, one might think that large-amplitude motions are highly correlated.[4]

In oblivion of such complexity one could treat the protein folding reaction in terms of theories of elementary chemical reactions where the native and unfolded states are in equilibrium, N⇆U, and the free energy change across the transition is

modeled by

$$\Delta G = \Delta G(\mathrm{H_2O}) - m_g D \quad (1)$$

where $\Delta G(\mathrm{H_2O})$ is the change in Gibbs free energy change in the aquoeous milieu, often also called conformational stability, m_g is the change in protein surface area across N and U states, and D is the concentrations of the chemical denaturant used to perturb the equilibrium.[5,6] Guanidinium chloride (GdnHCl) and urea are the universally adopted chemical denaturants to unfold the protein, the former being roughly three-times more effective. Because of the involvement of weak forces of interaction in protein structure, the unfolding process carried out by increments of a denaturant appears as a sharp cooperative transition, which is limited to provide values of K_U=U/N within the transition zone. A widely used procedure is to read out the value of $\Delta G(\mathrm{H_2O})$ by linearly extrapolating the ΔG dependence on D to the ordinate. Another linear extrapolation method for determination of $\Delta G(\mathrm{H_2O})$ is based on the details of protein groups that are solvent-exposed in the unfolded state and not in the native state.[7] Thus, in the equation,[6-9]

$$\Delta G = \Delta G(\mathrm{H_2O}) + \sum_i n_i \alpha_i \Delta \mathrm{g}_{\mathrm{tr},i} \quad (2)$$

n_i is the total number of i type of groups present in the protein, α_i represents the fraction of the i-type group exposed to the solvent in the U state, and $\Delta g_{\mathrm{tr},i}$measures the free energy of transfer of the i-type group from water to a given level of the chemical denaturant used. These linear free-energy relations (LFERs), especially equation 1, have been used to analyze equilibrium unfolding transitions of virtually all small single-domain proteins with the assumption that only the two states, N and U, exist across the equilibrium.

Literature reports many such two-state proteins whose folding reaction involves only the N and U states[10-12] allowing simple thermodynamic analyses using LFER. Some proteins, however small they may be, also involve one or more specific folding intermediates in the folding reaction.[13-16] Such multi-state proteins have been tabulated from time to time[12] Inclusion of intermediates and often complex deviation from the simple two-state behavior complicates the analysis of the folding reaction.

Clearly, all such classical models attach discreteness to the native, unfolded, and intermediate states, and the discreteness allows for physical description of the reaction in terms of phase behavior. Indeed, generic phase diagrams of discrete thermodynamic states of the folding reaction have been described.[17-21] By simple thoughts of structure and atoms packing in crystals vis a vis proteins, one is tempted to subscribe to these models, the phase behavior, and analyses.

However, the models and analysis procedures discussed above need reconsideration based on two issues regarding the interpretation of sigmoidal protein folding curves. One, the interpretation above assumes that the end states (N and U) are structurally invariant to denaturing variables so that the only information about folding is contained in the transition region of the curve defined as the region around the denaturation midpoint. Two, how valid is the linear free energy relationship (LFER) to describe the thermodynamics of protein folding as a function of the denaturing agent? Does the simplicity of LFER deserve enough merit for wide adaptability? Both these issues are tightly connected and have critical consequences to the understanding of the protein folding problem. Both are tightly roped in regard to accuracy, validity, and interpretation of thermodynamic parameters obtained from a two-state analysis of folding data, and both determine the extent to which important

changes that can take place during the unfolding reaction are ignored for the sake of simplicity.

The insufficiency of LFER can be judged from many reports of the past few years that have shown that the conformational stability, $\Delta G(H_2O)$, does not vary linearly with the concentrations of the denaturing agent, suggesting a breakdown of the two-state unfolding. Although a large part of these reports are based on fast-folding proteins, the nature of observations, the general implications for the analysis of folding data, and the interpretation of results in terms of non-classical multistate behavior are all similar irrespective of equilibrium and kinetic nature of studies. For example, it has been shown that the pre- and post-transition regions of sigmoidal unfolding curves can include large structural changes that are ignored in a two-state analysis.[22,23] Differential scanning calorimetric studies have shown that during the pre-transition unfolding region many proteins undergo large structural fluctuations and partial unfolding as determined by the heat capacity changes with temperature.[24,25] Cleverly designed experiments and improved sensitivity of monitoring probes can provide more information to understand the action of pre- and post-unfolding levels of denaturants on structure, dynamics, and energetics.

Singularity or discreteness of end states demand discrete high-energy barrier and relatively firm undeformable energy surfaces where transition from one state to another requires the crossing of at least one free energy barrier.[26] However, many fast folding experiments suggest that proteins pass through marginal barriers, and in some occasions transitions may occur with barrier size approaching zero. The protein is then said to fold globally downhill[27,28] which does not reconcile with specific-state folding models where states are separated by sizable energy barrier.

Models and interpretations must also include structural, dynamic, and energetic variation of the protein state in the post-transition region where the solvent quality improves substantially due to the large amount of denaturant(s) included in the medium.[22] Residual structure of the denaturant-specific unfolded state has been determined for a few proteins by NMR.[29-33] To this end, studies on dynamics, dimension, and energy of the post-transition protein state have been really scarce, mainly due to lack of well-defined structure and higher level of denaturant. Denaturant-dependent hydrodynamic properties of the unfolded protein chain for some proteins have been studied by single-molecule experiments,[34-36] and the results corroborate with the polymer theory-based argument of the similarity of global properties of homopolymers and denatured proteins.[26,37] The results provide evidence for continuous expansion of the denatured chain as the solvent quality turns increasingly better.[26]

This thesis thematically delves into dynamic and energetic responses of both N and U states to denaturant perturbation during the unfolding of three proteins - lysozyme, cytochrome *c*, and *At*PP1 (*Arabidopsisthaliana* phloem protein 1). The former two are paradigms for protein folding studies, and indeed numerous studies on these proteins have used LFER analysis for both treating equilibrium thermodynamic data, and describing a few kinetic intermediate(s) whose number depends on the solution conditions. *At*PP1 is a least studied protein. In fact, the atomic structure and folding of this protein are unknown, and no functional studies have been carried out to date. Based on known functions of closely related phloem protein homolog,[38] *At*pp1 is thought to transport nutrients and deliver mRNA to the phloem translocation system. Determination of the solution structure of *At*pp1 is in progress in this

laboratory, and the NMR resonance assignments required to carry out the work of this thesis have been completed already.

The properties of N-like states of lysozyme and *At*pp1 have been studied under subdenaturing conditions created by including lower levels of chemical denaturants where the protein does not denature globally. Global properties of U-like states are measured under conditions of high concentrations of GdnHCl limited by its aqueous solubility. Most of the experiments involve basic spectroscopic and NMR-based measurements, including ^{15}N relaxation, global translational diffusion, and temperature dependence of amide proton chemical shifts. Studies on the unfolded state were limited to cytochrome *c*, for which the method of pulsed laser-based photodissociation of heme-bound CO was employed. Each chapter of the thesis provides an introduction section, and is presented in a self-consistent manner with relevant citations.

1.2 References

1. Richards, F. M. The Interpretatopn ogf Protein Structures: Total Volume, Group Volume Distributions, and Packing Density.*J.Mol. Biol*.**1974**, *82*, 1-14.
2. Sarvazyan, A. P.; Kharakoz, D. P.; Hemmes, P. Ultrasonic Investigation of the pH-Dependent Solute-Solvent Interactions in Aqueous Solutions of Amino acids and Proteins. *J. Phys. Chem*. **1979**, *83*, 1796-1799.
3. Gekko, K.; Noguchi, H. Compressibility of Globular Proteins in Water at 25ºC. *J. Phys. Chem*. **1979**, *83*, 2706-2714.
4. Gavish, B.; Gratton, E.; Hardy, C. J. Adiabatic Compressibility of Globular Proteins. *Proc. Natl. Sci. Acad.USA***1983**, *80*,750-754.
5. Greene, R. F.; Pace, C. N. Urea and Guanidine Hydrochloride Denaturation of Ribonuclease, Lysozyme, α-Chymotrypsin, and β-Lactoglobulin.*J. Biol. Chem*. **1974**, *249*, 5388-5393.
6. Pace, C. N.; Shaw, K. L. Linear Extrapolation Method of Analyzing Solvent Denaturation Curves. *Proteins***2000**, *4*, 1-7.
7. Tanford, C. Isotermal Unfolding of Globular Peoteins in Aqueous Urea Solutions. *J. Am. Chem. Soc*. **1964**, *86*, 2050-2059.
8. Tanford, C. Protein Denaturation: Part C. Theoretical Models for the Mechanism of Denaturation.*Adv. ProteinChem*. **1969**, *24*, 1-95.
9. Jackson, S. E.; Fersht, A. R. Folding of Chymotrypsin Inhibitor 2. 1. Evidence for a Two-State Transition.*Biochemistry***1991**, *30*, 10428-10435.
10. Jackson, S. E. How do Small Single-Domain Proteins Fold? Folding *Des*. **1998**, *3*, 81-91.
11. Prabhu, N. P.; Bhuyan, A. K. Prediction of Folding Rates of Small Proteins:

Empirical Relations Based on Length, Secondary Structure Content, Residue Type, and Stability. *Biochemistry***2006**, *45*, 3805-3812.

12. Rollins, G. C.; Dill, K. A. General Mechanism of Two-State Proteins Folding Kinetics. *J*. *Am*. *Chem*. *Soc*. **2014**, *136*, 11420-11427.
13. Kim, P. S.; Baldwin, R. L. Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Annu*. *Rev*. *Biochem*. **1982**, *51*, 459-489.
14. Kim, P. S.; Baldwin, R. L. Intermediates in the Folding Reactions of Small Proteins. *Annu*. *Rev*. *Biochem*. **1990**, *59*, 631-660.
15. Ptitsyn, O. B. Kinetic and Equilibrium Intemediates in Protein Folding. *ProteinEngineering***1994**, *7*, 593-596.
16. Baldwin, R. L. The Nature of Protein Folding Pathways: The Classical versus the New View. *J*. *Biomol*. *NMR***1995**, *5*, 103-109.
17. Finkelstein, A. V.; Shakhnovich, E. I. Theory of Cooperative Transitions in Protein Molecules. II. Phase Diagram for a Protein Molecule in Solution. *Biopolymers***1989***28*, 1681-1694.
18. Pande, V. S.; Roshkar, D. S. Is the Molten Globule a Third Phase of Proteins? *Proc*. *Natl*. *Sci*. *Acad*. *USA***1998**, *95*, 1490-1494.
19. Kamatari, Y. O.; Ohji, S.; Konno, T.; Seki, Y.; Soda, K.; Kataoka, M.; Akasaka, K. The Compact and Extended Denatured Conformation of Apomyoglobin in the Methanol-Water Solvent. *ProteinSci*. **1999**, *8*, 873-882.
20. Hermans, J.; Lentz, B. *EquilibriaandKineticsofBiological Macromolecules*; John Wiley & Sons: Hoboken, New Jersey, **2014**.
21. Atkins, P.; de Paula, J. *PhysicalCemistryfortheLifeScinces*; Oxford University

Press, Oxford, **2011**.

22. Naganathan, A. N.; Perez-Jimenes, R.; Sanchez-Ruiz, J. M.; Munõz, V. Robustness of Downhill Folding: Guidelines for the Analysis of Equilibrium Folding Experiments on Small Proteins. *Biochemistry***2005**, *44*, 7435-7449.
23. Oliva, F. Y.; Munõz, V. A Simple Thermodynamic Test to Discriminate Between Two-state and Downhill Folding. *J. Am. Chem. Soc.* **2004**, *126*, 8596-8597.
24. Munõz, V.; Sanchez-Ruiz, J. M. Exploring Protein Folding Ensembles: A Variable-barrier model for the Analysis of Equilibrium Unfolding Experiments. *Proc. Natl. Acad. Sci. USA***2004**, *101*, 17646-17651.
25. Naganathan, A. N.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Munõz, V. Estimation of Protein Folding Free Energy Barriers from Calorimetric Data by Multi-model Bayesian Analysis. *Phys. Chem. Chem. Phys.***2011**, *13*, 17064-17076.
26. Dill, K. A.; Shortle, D. Denatured States of Proteins. *Annu. Rev. Biochem.* **1991**, *60*, 795−825.
27. Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Munõz, V. Experimental Identification of Downhill Protein Folding. *Science*,**2002**, *298*, 2191−2195.
28. Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels. *J. Chem. Phys*, **1996**,*104*, 5860.
29. Neri, D.; Wider, G.; Wuthrich, K. Complete ^{15}N and ^{1}H NMR Assignments for the Amino Terminal Domain of the 434 Repressor in the Urea-Unfolded Form. *Proc. Natl. Acad. Sci. USA***1992**, *89*, 4397-4401.
30. Neri, D.; Billeter, M.; Wider, G.; Wuthrich, K. Structure of the Urea-Unfolded 434 repressor Determined by NMR in Solution. *Science***1992**, *257*, 1559-1563.

31. Zhang, O.; Kay, L. E.; Olivier, J. P.; Forman-Kay, J. D. Backbone 1H and ^{15}N Resonance Assignments of the N-Terminal SH3 Domain of drk in Folded and Unfolded States Using Enhanced-Sensitivity Pulsed Field Gradient NMR Techniques. *J. Biomol. NMR***1994**, *4*, 845-858.
32. Chatterjee, A.; Kumar, A.; Chugh, J.; Srivastava, S.; Bhavesh, N. S.; Hosur, R. V. NMR of Unfolded Proteins. *J. Chem. Sci.* **2005**, *117*, 3-21.
33. Kosol, S.; Contreras-Martos, S.; Cedeno, C.; Tompa, P. Structural Characterization of Intrinsically Disordered Proteins by NMR Spectroscopy. *Molecules***2013**, *18*, 10802-10828.
34. Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the free energy surface for protein folding with single molecule fluorescence spectroscopy. *Nature***2002**, *419*, 743-747.
35. Ziv, G.; Haran, G. Protein folding, protein collapse, and Tanford's transfer model: lessons from single molecule FRET. *J. Am. Chem. Soc.* **2009**, *131*, 2942-2947.
36. England, J. L.; Haran, G. Role of salvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu. Rev. Phys. Chem.* **2011**, *62*, 257-277.
37. Dill, K. A. Theory for the folding and stability of globular proteins. *Biochemistry***1985**, *24*, 1501-1509.
38. Xoconostle-Cazares, B.; Xiang, Y.; Ruiz-Medrano, R.; Wang, H. L.; Monzer, J.; Yoo, B. C.; McFarland, K. C.; Franceschi, V. R.; Lucas, W. J. Plant Paralog to Viral Movement Protein that Potentiates Transport of mRNA into the Phloem. *Science***1999**, *283*, 94-98.

Chapter-2

Free Energy Landscape of Lysozyme: Multiple Near-Native Conformational States and Rollover in the Urea Dependence of Folding Energy.

2.1 Abstract

Deviation from linearity of the equilibrium folding free energy (ΔG) of proteins along the reaction coordinate is scarcely known. Optical spectroscopic observables and NMR-measured average molecular dimensional property of lysozyme with urea at pH 5 reveal that ΔG rolls over from linearity under mild to strongly native-like conditions. The urea dependence of ΔG is graphed in the 0-7 M range of the denaturant by employing a series of guanidine hydrochloride (GdnHCl)-induced equilibrium unfolding transitions, each in the presence of a fixed level of urea. The observed linear dependence of ΔG on urea under denaturing conditions begins to deviate as moderately native-like conditions are approached, and eventually rolls over under strongly native-like conditions. This is atypical of the upward curvature in the ΔG vs denaturant plot predicted by the denaturant binding model. On increasing the denaturant concentration from 0 to 5 M, the hydrodynamic radius of lysozyme shrinks by ~2 Å. We suggest subdenaturing levels of urea affect the population distribution amongst multiple near-native isoenergetic conformational states so as to promote them sequentially with increments of the denaturant. We use a multiple-state sequential model to show that the keel over of ΔG occurs due to these near-native alternative states in the native ensemble used for defining the unfolding

equilibrium constant (K_U), which we assume to vary linearly with urea. The results and the model appear to indicate a rugged flat bottom in the free energy landscape wherein population distribution of native-like states is modulated by urea-affected interstate motions.

2.2 Introduction

Naively, it is easy to understand the protein folding equilibrium transition. If atom packing in a typical native-state protein molecule is likened to the atom packing in a metal or crystal-like solid substance, then the melting of the solid protein should be described by a highly cooperative biphasic transition where the initial and final states, called end states, are individually invariant to the temperature or the chemical denaturant employed to induce the transition. The thought gains footage from the compilation of a sizable set of proteins that appear to fold-unfold, as judged from published reports, in a biphasic two-state manner.[1-6] Some relatively recent studies however state that the two-state unfolding process for many proteins could just be an appearance. Closer examination often reveals structural and dimensional changes in both end states in response to denaturing agents, suggesting a breakdown of the two-state approximation. Denaturant-dependent changes in end states in parallel to their two-state-like interconversion in the equilibrium unfolding of the slow folder alpha-spectrin SH3 have been reported recently by NMR and single-molecule spectroscopy.[7] Large structural changes in the pre- and post-transition regions of the sigmoidal unfolding curve of the hitherto acclaimed two-state folder BBL,[8] and partial unfolding and structural fluctuations in the pre-transition region as determined by the heat capacity changes with temperature have been shown for some

proteins.[9,10]The deviation from the two-state approximation for BBL unfolding is also observed in experiments that employ both chemical denaturant and temperature.[11]

Different unfolding curves depending on the spectroscopic probe can provide more direct evidence for non-compliance with two-state unfolding approximation.[8,12-15] However, multiple measurement probes applied to small proteins most often report on the same unfolding transition providing no evidence for intermediate population, and data are then simply analyzed by a two-state model ($N \rightleftharpoons U$) where the free energy difference between the U and N state ensembles (ΔG) is assumed to vary linearly with denaturant concentration,[16] although other assumptions of the free energy-denaturant relationship remain open to scrutiny.[17-19]The basic insufficiency of such data arises from the general silence or poor response of common spectroscopic probes, including fluorescence, CD, visible optical, and UV difference absorbance, to changes in the properties of native and unfolded ensembles outside the global transition region. The probes report on the denaturant-induced changes in the proportion of U and N end states from which the ΔG dependence on the denaturant is conveniently graphed within the narrow cooperative transition region of the unfolding curve. Values of ΔG in the pretransition region can be obtained from a series of extended experiments in which unfolding is carried out isothermally by using another denaturant while holding the test denaturant level fixed at concentrations within the pretransition region.[17,18,20,21] Such analyses reveal not only the functional dependence of ΔG on denaturant, but could also provide insight of native-state ensemble properties.

Occasionally denaturant response of structural and thermodynamic properties of an end-state ensemble under favorable conditions of pH and temperature can be observed by one or more of the common spectroscopic methods, or by a careful

choice of a global observable.[19] Such intra-state data may not provide information on the functional dependence of ΔG on denaturant per se, but are invaluable to learning about structural intermediates, conformational hierarchy, and free energy landscape. For example, the occurrence of a near-native equilibrium intermediate of the protein G B1 domain was observed by the record of the global observable Stokes radius,[19] partially unfolded forms (PUFs) of cytochrome *c* were uncovered by native-state hydrogen exchange,[22,23] and continuous expansion of the cytochrome *c* chain within the unfolded state has been reported by probing the ensemble average hydrodynamic radius with increasingly unfolding concentration of the denaturant.[24]

Both of the above, changes within the native state and mapping of ΔG under near-native conditions, are studied here with hen lysozyme at pH 5, 25°C. Folding studies of lysozyme are generally carried out by using GdnHCl as the denaturant, and since the common spectroscopic observables do not report on changes in the pretransition region of the equilibrium unfolding curve, data are analyzed by the two-state N⇌U transition,[18,25,26] even though solution X-ray scattering studies of urea-induced unfolding at pH 2.9 appeared to reveal an equilibrium intermediate.[27]Unfolding of lysozyme is largely incomplete within the solubility limit of urea in the pH 4-9 region. However, fluorescence and near-UV CD signals and associated spectral parameters undergo pronounced changes even in subdenaturing concentrations of urea. These changes are accompanied by shrinkage in the molecular dimension and substantial loss of lysozyme enzyme activity. The urea dependence of ΔG, which is linear under denaturing conditions, rolls over as native-like conditions are approached. These results are explained by invoking changes in the native state ensemble so that alternative near-native states are populated with increments of urea.

The model retains the linear free energy-denaturant formalism, but expresses the unfolding equilibrium constant (K_U) in terms of the U state and a large number native-like isoenergetic conformational states. The data and the model are discussed in the perspective of multiple conformational states of the native free energy landscape.[28]An alternative simpler model based on one-dimensional free energy surface approach[6,29] is also cited.

2.3 Experimental Section

2.3.1 Equilibrium Unfolding. Samples for equilibrium unfolding measurement were prepared by mixing appropriate volumes of the native and the unfolded protein stock solutions. The two stock solutions were identical in protein concentration, which was 5, 10, and 100 μM for fluorescence, near-UV, and far-UV measurements, respectively. The stock solutions were buffered in 20 mM sodium acetate, pH 5. In experiments involving titration with GdnHCl holding urea concentration constant, the two stock solutions were made identical in terms of urea content as well. The mixed samples were incubated at 25°C for about 12 h before fluorescence and CD measurements. Emission spectra were recorded in the 320-375 nm region by exciting at 280 nm in a Fluoromax 4P (Horiba) instrument. CD spectra in the 290 and 225 nm regions were recorded using a AVIV 420SF spectrometer. Denaturant concentrations in samples were determined from their refractive index readings.

2.3.2 Enzyme Kinetics. The rate of lysozyme-catalyzed hydrolysis of $(GlcNAc)_3$-MeU (4-methylumbelliferyl β-D-N,N′,N″-acetylchitotroside) in the

presence of different concentrations of urea was measured following the procedure of Yang and Hamaguchi.[30] Briefly, the reaction mixture consisted of 0.14 μM lysozyme, 12.73 μM $(GlcNAc)_3$-MeU, and a given amount of urea in 20 mM acetate, pH 5.3. After mixing lysozyme with the substrate at 41.5°C, aliquotes of the mixture were withdrawn at various times in the 0-80 min regime, and the reaction was stopped by adding 50 mM glycine, pH 13. The cleavage product, 4-methylumbelliferone, in its negatively charged state was quantified by fluorescence emission at 450 nm (excitation, 360 nm). As described earlier,[30] a linear calibration curve was used to determine the amount of 4-methylumbelliferone formed. Experimental values of the product formed were corrected for the effect of urea in control experiments that excluded lysozyme. Initial velocity of the reaction was determined from the initial linear region of the product-time graph.

2.3.3 NMR Measurements. All experiments were done using 2 mM lysozyme in 10 mM sodium acetate, pH 5, 25°C using a 500 MHz Avance III Bruker spectrometer. Deuterated urea was prepared by repeated drying of urea from D_2O solutions under pure nitrogen gas. Hydrodynamic radius (R_H) of lysozyme in the presence of variable amount of urea was determined using the sLED-PFG pulse sequence[31] with diffusion gradient (*z*-gradient) strength in the range of 3-50 Gauss cm^{-1}. The samples also contained 0.5 mM 1,4-dioxane for internal R_H standard.[32] Two-dimensional experiments were performed using the basic sequences. Values of R_H were calculated by

$$\mathrm{I(g)} = A \exp(-\mathrm{kg}^2) \qquad (1)$$

$$R_H^{protein} = R_H^{dioxane}\left(\frac{k_{dioxane}}{k_{protein}}\right) \quad (2)$$

where, *I* and g are NMR signal intensity and gradient strength, respectively.

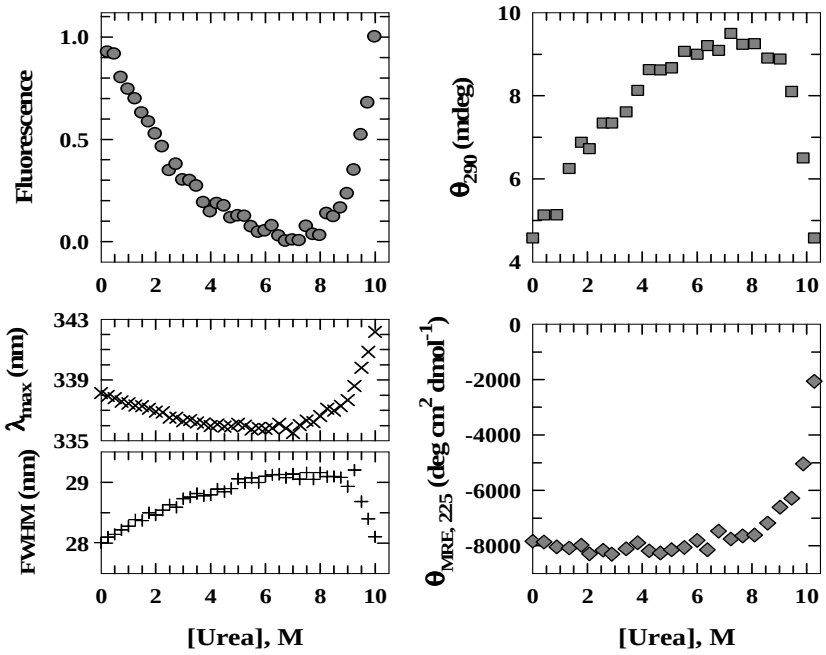


Figure 1. Changes in common spectroscopic observables of lysozyme with urea in 20 mM sodium acetate, pH 5, and 25°C. (a) Tryptophan fluorescence emission is increasingly quenched with denaturant increment in the pretransition (subdenaturing) region, and is entirely quenched before global unfolding sets in. (b) The wavelength maximum of fluorescence emission blue shifts by ~ 3 nm in the subdenaturing region, but shifts largely to longer wavelength in the transition region. (c) Inhomogeneous broadening of the emission band in the subdenaturing regions. (d) Accentuated near-UV CD signal in the pretransition region also indicates tertiary structural changes before global unfolding. (e) Far-UV CD provides no indication of secondary structural changes at subdenaturing concentrations of urea.

2.4 Results

2.4.1 Tertiary Structural Changes in Subdenatured Lysozyme in Urea at pH 5. The large body of lysozyme folding studies has generally employed GdnHCl due to inability of urea to disrupt the structure adequately. Indeed, lysozyme remains largely folded within the aqueous solubility limit of urea, especially in the pH 4-9 zone (Supporting Information, SI Figure 1). At pH 5, the experimental condition employed in this study, the urea dependence of the quantum yield of tryptophan fluorescence is quenched almost completely as the denaturant level is raised up to ~7 M, but increases thereafter (Figure 1a). Similarly, the emission wavelength maximum decreases systematically from 338 nm by ~3 nm in the 0-7 M range of urea, but increases with higher urea (Figure 1b). This wavelength shift probably originates from greater burial of the fluorophores in the low dielectric protein interior. Concomitantly, the FWHM of the Gaussian emission band increases by ~1 nm up to ~7 M urea, and decreases at higher concentrations (Figure 1c). The increase in the inhomogeneous spectral broadening reflects increased collisional quenching of fluorophores, the distribution of whose small changes in the urea-dependent environment shifts with increasing denaturing conditions. These results suggest changes in the protein tertiary structure up to ~7 M urea followed by large-scale structure unfolding. The conjecture is supported by CD results. The large increase in the near-UV CD signal in the 0-7 M range of urea (Figure 1d) indicatescontinuous changes in geometry and decrease in interchromophore distances so as to augment rotational strengths of aromatic side chains by increased coupling of theirelectric transition dipole moments - an explanation based on oscillator coupling mechanism.[33] On the other hand, urea-dependent changes in the far-UV CD signalindicates

disruption of secondary structures only when the denaturant level is raised above 7 M (Figure 1e). Overall, urea induced unfolding of lysozyme at pH 5 seems to involve tertiary structural changes before the major cooperative transition sets in. The nature and magnitude of changes in spectral parameters appear to hint at molecular contraction of native lysozyme in the denaturing limit of urea.The basis of spectroscopic probe-dependent unfolding curves,[8,13-15] can often be used to perform an atomic resolution analysis of equilibrium unfolding processes of fast-folding proteins using NMR.[34]

2.4.2 Urea Dependent Hydrodynamic Radius of Lysozyme. To test the possibility of shrinkage of the average molecular dimension in denaturing concentrations of urea, the average hydrodynamic radius, $<R_H>$, of lysozyme was determined by pulsed field gradient NMR. The $<R_H>$ decreases from ~21.5 to 19.6 Å in the 0-6 M range of urea and increases at higher urea level (Figure 2). While $<R_H>$ is expected to increase at higher urea due to large-scale unfolding of the protein, the initial decrease of ~2 Å, observed by others as well,[32] suggests a contraction of molecular dimension as it is placed in increasingly denaturing conditions. It thus appears that subdenaturing levels of urea at pH 5 not only altersthe tertiary structure of lysozyme, but also causes molecular dimensional shrinkage.

2.4.3 Enzyme Activity of Lysozyme Under Denaturing Conditions. Because enzyme activity is sensitive to molecular conformation, lysozyme-catalyzed hydrolysis of the synthetic substrate $(GlcNAc)_3$-MeU was determined in the presence of urea. The initial velocity of the reaction decreases gradually and is reduced by

~60% of the native state before global unfolding sets in. This observation further indicates gradual alteration of lysozyme conformation with increasing denaturing conditions.

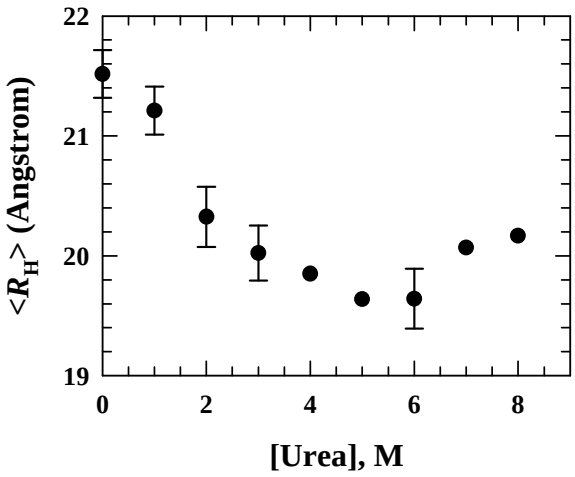


Figure 2. NMR-determined average hydrodynamic radius of lysozyme in 10 mM sodium acetate, pH 5, 25°C decreases by ~2 Å in going from 0 to 6 M urea, but increases upon approaching the unfolding transition region. Error bars are drawn from multiple repeats of the experiment.

2.4.4 Absence of Equilibrium Intermediates. None of the spectroscopic measurements however suggest distinct populations of equilibrium intermediates. The examination of a series of NOESY spectra across the subdenaturing region of urea concentration (0-7 M) does show some minor changes in peak intensities and distance constraints (Supporting Information, SI Figure 2), but they may not be adequate to draw any conclusion about equilibrium intermediate populations. They nonetheless are sufficient to allow for the inference of progressive changes in tertiary structure under increasing subdenaturing conditions.

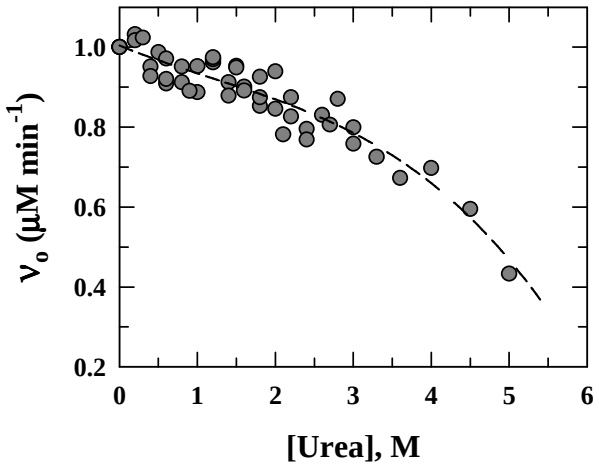


Figure 3. Initial velocity of lysozyme-catalyzed hydrolysis of the model substrate $(GlcNAc)_3$-MeU at subdenaturing concentrations of urea in 20 mM sodium acetate buffer, pH 5.3, 41.5°C. The velocity values are shown relative to that for native lysozyme. The dashed line through data has no physical basis.

2.4.5 GdnHCl-Induced Unfolding of Urea-Denatured Lysozyme at pH 5. To further understand the process of lysozyme denaturation we performed a series of experiments where the protein denatured to variable extent in the presence of low concentrations of urea at pH 5 was titrated by GdnHCl, and unfolding was monitored by fluorescence and far-UV CD (Supporting Information, SI Figure 3). Because no evidence for equilibrium intermediate of lysozyme has been found here and reported in earlier studies,[18,25] the unfolding data were analyzed by the two-state model to extract the urea dependence of the folding free-energy, ΔG (Figure 4). Strictly, these ΔG values need correction for the effect of urea on GdnHCl unfolding of lysozyme.[18]The correction could not beintroduced here because of ineffectiveness of urea to completelyunfold theprotein at pH 5 even in the presence of low concentrations of

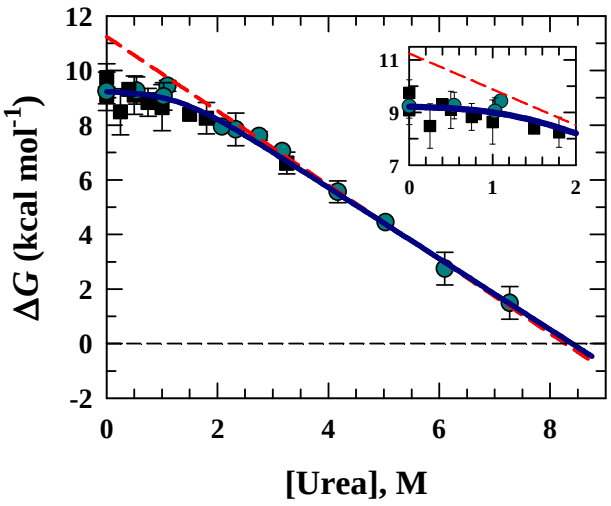


Figure 4. Free energy of lysozyme folding as a function of urea in 20 mM sodium acetate, pH 5, 25°C. Values of Δ*G* plotted were extracted from two-state analysis of GdnHCl denaturation curves, each for a different concentration of urea. Monitoring probes were fluorescence (■) and far-UV CD (●). The curved solid line (blue) is the best fit of the data according to equation 3 of the sequential near-native states of the model. The straight dashed line (red) is according to equation 4, which accounts for a single native state instead of multiple near-native states. The rollover region of the plot is enlarged in the *inset*.

GdnHCl, and we assume no significant effect would be produced without the correction. Figure 4 shows that Δ*G* conspicuously rolls over at lower concentrations of urea. The deviation is reproducible irrespective of the monitoring probe, and is outside the estimated error limit. This result provides evidence that urea dependence of Δ*G* is inconsistent with the linear free-energy model. Rollover in the Δ*G*-denaturant isotherm should imply the presence of native-like isoenergetic species across the folding-unfolding equilibrium. In thermodynamic terms, different states of the protein are constituted by different numbers of denaturant binding. Although no definitive indication of stable intermediate populations have been found, we utilize the evidences for progressive changes in tertiary structure and $\langle R_H \rangle$ under

increasingly denaturing conditions to suggest the existence of several rapidly interconverting conformations that have tertiary structures destabilized to variable extent but native-specific chain topology. With respect to these species under subdenaturing conditions, the urea-induced unfolding process of lysozyme at pH 5 is given by the following equilibria

$$\mathrm{N} \overset{K_1}{\rightleftharpoons} \mathrm{N_1}^* \overset{K_2}{\rightleftharpoons} \mathrm{N_2}^* \overset{K_3}{\rightleftharpoons} \cdots\cdots \rightleftharpoons \mathrm{U}$$

where, N^*_i are native-like denatured states. We define an apparent equilibrium constant, K_{app}, which assumes more significance under subdenaturing conditions where the protein ensemble is strongly native-like. At low concentrations of urea,

$$K_{\mathrm{app}} = K_1 + K_1K_2 + K_1K_2K_3 + \cdots + K_{\mathrm{U}}(1 + K_1 + K_1K_2 + K_1K_2K_3 + \cdots) \qquad (3)$$

where, the equilibrium constant for global unfolding, K_{U}, is given by

$$K_{\mathrm{U}} = \frac{K_1K_2K_3 \ldots}{1 + K_1 + K_1K_2 + K_1K_2K_3 + \cdots} \qquad (4)$$

Using this expression for K_{U} in equation 3, the free-energy of unfolding, ΔG, can be written as

$$\Delta G = -RT\ln\left\{a + K_U^0 e^{\frac{m_{\mathrm{g}}[\mathrm{urea}]}{RT}}(1 + a)\right\} \qquad (5)$$

where, $a = K_1 + K_1K_2 + \cdots + (K_1K_2 \cdots K_{n-1})$ for n equilibrium constants corresponding to $n+1$ states, and $K_{\mathrm{U}} = \mathrm{K_U}^{\circ} \exp(m_{\mathrm{g}}[\mathrm{urea}]/\mathrm{RT})$, which is the usual linear free energy relation.

Equation 3 simulates the observed nonlinear dependence of ΔG at low urea (Figure 4). Under subdenaturing to native-like conditions, the lower order equilibrium constants are little dependent on the denaturant level. As strongly denaturing

conditions are approached, the value of K_U turns largely dependent on urea concentration, and is much larger than those of the equilibrium constants, K_i. Equation 5 then reduces to

$$\Delta G = -RT\ln\left\{K_U^0 e^{\frac{m_g[\text{urea}]}{RT}}\right\} \quad (6)$$

The results and the model thus imply a rough free energy surface under native-like or subdenaturing conditions consisting of a large number of interconverting conformational states, which share native chain topology and secondary structures, but are different from each other in terms of tertiary structure, side-chain geometry, and average molecular dimension. In the context, the two-denaturant (urea and GdnHCl) experiment performed here is similar to the double-perturbation unfolding (denaturant and temperature) of the protein BBL which also showed non-linearity in the urea dependence of temperature parameters.[11]

2.5 Discussion

The native-state ensemble of a protein consists of multiple conformational states, which are often not sensed by common spectroscopic methods that generally report on the relative population of end states in the transition region of the unfolding equilibrium, N⇌U. Although such results provide the denaturant dependence of free energy reasonably correctly for most proteins, the conformational and global properties of end-state ensembles remain largely unknown. Under favorable conditions, however, denaturant-induced changes in end-state properties can be measured by ensemble average experiments using specific local and global probes. Reports of GdnHCl-induced molecular expansion of the native-state ensemble of

protein β1,[19] and the unfolded ensemble of horse cytochrome c[24] provide earlier examples of such measurements. For lysozyme at pH 5 here, both fluorescence and near-UV CD signals are highly sensitive to properties of the native-like ensemble which characterizes the pretransition region in urea-induced denaturation (Figure 1). It is also significant that the pretransition region is associated with a gradual shrinkage of the ensemble average molecular dimension by ~2 Å (Figure 2), consistent with our earlier observation with relaxation dispersion NMR spectroscopy.[35] These changes expectedly abrogate the enzyme activity substantially (Figure 3). It emerges that conformational, thermodynamic and functional properties of proteins can undergo dramatic changes even before global chain unfolding ensues.

The variable properties of the native-like ensemble must be taken into account in the determination of the free energy difference between the end states at various points across the N⇌U equilibrium. Generally, the graphed denaturant dependence of ΔG in the cooperative transition region is linearly extrapolated toward the pretransition region (linear extrapolation method, LEM) to obtain the free energy difference between the unfolded and native states, ΔG.[16-19,36] The LEM is particularly suited for graphing ΔG dependence on urea because the alternative binding site model[37] appears fraught with several problems.[18] The analysis presented here (equations 2-4, Figure 4) indeed assumes linear functional dependence of the unfolding equilibrium constant, K_U, on urea. However, K_U in the model includes not only the N and U end states, but also the ensemble of native-like conformational states (N_i^*) whose properties and relative population are sensitive to subdenaturing concentrations of urea. It is this inclusion of K_i^* in the expression for K_U which produces the rollover in the free energy-denaturant graph. The distinct conformational

states of the native-like ensemble disappear in the transition region where K_U is defined by the relative population of end states N and U alone, and then ΔG varies linearly with urea. Thus, variable plasticity, and conformational and thermodynamic properties of the native-like ensemble can produce the downward curvature in the plot of denaturant dependence of folding free energy.

The free energy roll over is not just an oddity limited to equilibrium unfolding results for lysozyme here. A large body of work reported in the recent past has consistently provided evidence for rollover of the folding limb of kinetic chevron plots when the proper time-resolution is achieved to resolve the burst-phase of folding intermediates. Such rollovers need not always imply the presence of free-energy barriers,[38] because the folding reaction may be limited by not necessarily barriers but very low diffusion rates[39] or longer chain reconfiguration time under native-like folding conditions.[40] Folding limb rollover without significant energy barrier implies that the denaturant dependence of free energy is not linear in general terms.

Regarding the native-like ensemble model proposed here, a discussion on the details of the near-native conformational states (N_i^*) and their distribution is beyond the scope due to complexity and hyperdimensionality of protein conformational spaces. The results do however point to some isoenergeticity of the native-like states, often also called substates,[28] whose distributions are modulated continuously under variable subdenaturing conditions. The similarity of energies of these states produces some flatness in the overly simplified two-dimensional free energy surface under native-like conditions (Figure 5). If these conformational states constitute a corrugated harmonic surface where the state population is modulated by the presence of urea, the energy needed for their conversion to the unfolded state will be higher

relative to the threshold energy of conformational unfolding from an anharmonic surface. Such a free energy surface of singly-hierarchical modes has been described in the theoretical investigation of jumping-among-minima (JAM) model of energy landscape of native lysozyme, where transitions between conformational states may arise from fluctuations in side chain dihedral angles.[41] We assume that these fluctuations are widespread, and are affected by protein-urea interactions at subdenaturing levels of the denaturant, akin to the modulation of population distribution in the microstate ensemble of staphylococcal nuclease caused by proton binding.[42] Such fluctuations can be thermodynamically treated as local unfolding processes.[42] Because the features of the energy landscape change from one condition to the other, the distribution and interstate transition dynamics in the case of lysozyme are also expected to vary with different pH and denaturant. Associated with these transitions is a continuous decrease in hydrodynamic radius (Figure 2) the origin of which cannot be ascertained at present, but molecular dynamics simulation studies have also indicated that gyration radius is an important global parameter for description of energy landscape and conformational states of native proteins.[43,44] These global changes require large-scale motion of backbone C_α atoms, but we have obtained little evidence for collective motion. Nonetheless, as depicted in the free energy surface as a function of unfolding reaction coordinate (Figure 5), the measured folding free energy reports on the free energy difference between the end states U and N if subdenaturing levels of the denaturant produce little change in properties and population distribution of the N-state ensemble. When several near-native states are present and the interstate dynamics and population distribution are influenced by subdenaturing levels of the denaturant, unfolding measurements would sample the

free energy difference between the U state and the N^*-state ensemble along the reaction coordinate.

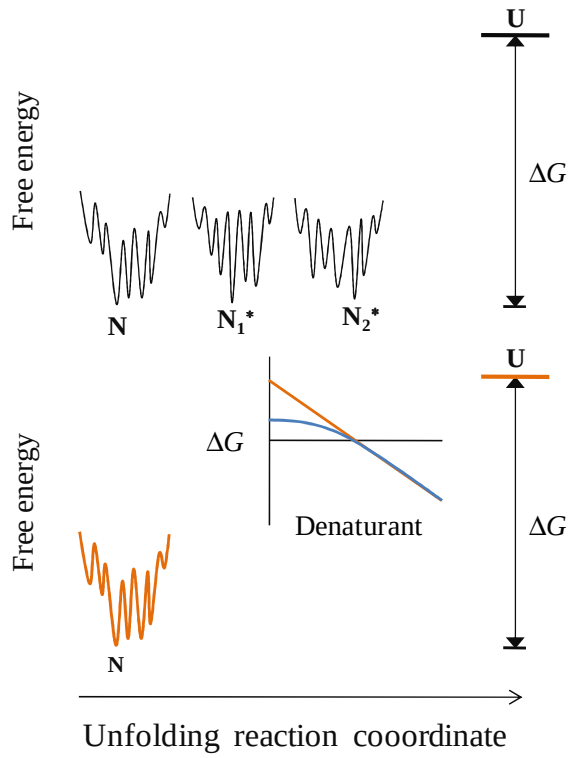


Figure 5. Highly simplified schematic of free energy-reaction coordinate space for lysozyme unfolding. The upper part of the figure shows the rugged harmonic surface along the reaction coordinate along which N_i^* states in equilibrium are sequentially populated as depicted by the model in the text. The N_i^* conformational states are like N in terms of free energy, and interstate transition rates are affected by subdenaturing concentrations of the denaturant. The ΔG value measured for such a system would be the difference in free energies of the unfolded state and the N_i^* state populated at the given denaturant concentration (the blue curved line in the center panel). The lower part of the figure depicts the expectation when N_i^* states are not populated or when the energy surface is not rugged. The measured ΔG value in this case will be equal to the difference in free energies of U and N states (the orange straight line in the center panel).

The complexity of lysozyme energy landscape possibly consisting of many high-energy structural intermediates has been surmised earlier.[26] Under favorable conditions these hidden intermediates should be detectable by native state hydrogen

exchange (HX) and analogous methods,[22,23,45-47] but pulse-labeling HX studies on lysozyme have not revealed much other than the operation of parallel folding pathways[48,49] and the population of a near-native state on the slow refolding pathway,[50] implying that lysozyme landscape probably consists of more near-native low-energy conformations than high-energy intermediates. Unlike the detection of high-energy equilibrium intermediates by HX-based methods, probing dynamical nature of near-native conformational states require other approaches, including relaxation dispersion-based NMR,[51] and Mossbauer and neutron scattering strategies.[52] While the discussion here invokes the multiple conformational state model to interpret the rollover in the free energy-denaturant space, dynamical details of the ensemble and the catalytic function of lysozyme under subdenaturing conditions are open to investigation.

An alternative model that can explain the effect of native-state structural change on the folding barrier is due to Munõz and coworkers.[6,29] The model is based on one-dimensional free energy surface for folding, and was initially developed to respond to the apparent but unrealistic two-state folding of diffusion-limited downhill folders.[29] Here, the size of the folding barrier is determined by the exponent of the stabilization energy. The model assumes that the order parameter for nativeness (scaled from 0 to 1 to denote completely unfolded and completely native structure, respectively) can serve as a reaction coordinate. Under this assumption the folding barriers are fully determined by the free energy surface, and the thermodynamic and kinetic barriers are indistinguishable.[29] The interpretation, applied recently to the folding of α-spectrin SH3 domain, assigns the denaturant-dependent structural changes in the native and unfolded states to the deformation of the inherent folding

free energy landscape.[7] These deformations of the folding free energy landscape are expected from a purely physical standpoint, and hence could be quite general (Figure 6). The perturbation in the magnitude of the folding barrier as a function of stabilization energy and other aspects of the model have been studied theoretically.[29] The results obtained in this study do not allow one to distinguish between the conformational substate model and the surface deformation model.

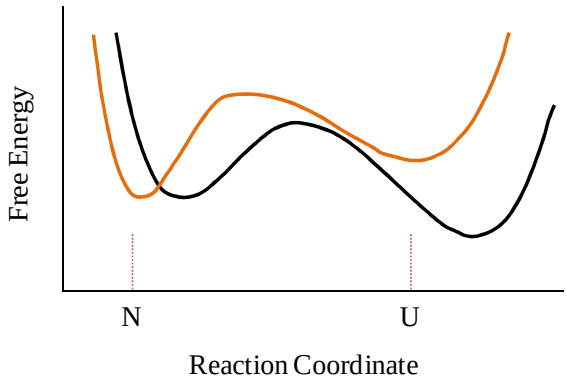


Figure 6. Deformations and shifts in the one-dimensional free energy surface when the N and U states under two-state approximation (red curve) become sensitive, both structurally and dimensionally, to the variation in denaturant level (black curve). The surfaces are redrawn after Campos et al.[7]

The ensemble-average experiments of the present study have focused on dimensional, structural, conformational, and energetic changes under native-like conditions. But equally important are responses of the denatured state to different levels of the denaturing agent. Expansion of the unfolded protein chain and associated internal frictional properties under variable denaturing conditions have been studied in sufficient details during the past decade or so by using single-molecule spectroscopy,[7,24,53-61] and to a lesser extent by ensemble-average experiments.[24] The compactness and dimension of the unfolded state under native-like conditions will affect folding thermodynamics,becausethese parameters determine the chain-

configurational entropy of the unfolded state,and hence the relative energies of the end states. Diminishing configurational entropy of the unfolded state with increasingly native-like conditions will obviously produce a roll over in the ΔG vs denaturant plot. We thus see that the denaturant-dependent variability of not only the native state but also the unfolded state of lysozyme could be the reason for the observed nonlinearity in the ΔG vs urea graph (Figure 4). To this end, the denatured state of lysozyme remains to be studied in detail so as to learn about its contribution to the free energy roll over.

2.6 References

1. Jackson, S. E. How do Small Single-domain Proteins Fold? *FoldingDes*. **1998**, *3*, 81-91.
2. Baldwin, R. L.; Rose, G. D. Is Protein Folding Hierarchic? II. Folding Intermediates and Transition States. *TrendsBiochem*. *Sci*.**1999**, *24*, 77-83.
3. Maxwell, K. L.; Wildes, D.; Zarrine-Afsar, A.; de Lod Rios, M. A.; Brown, A. G.; Friel, C. T.; Hedberg, L.; Horng, J-C.; Bona, D.; Miller, E. J.; Vallee-Belisle, A.; Main, E. R. G.; Bemporad, F.; Qiu, L.; Teilum, K.; Vu, N-D.; Edwards, A. M.; Ruczinski, I.; Poulsen, F. M.; Kragelund, B. B.; Michnick, S. W.; Chiti, F.; Bai, Y.; Hagen, S. J.; Serrano, L.; Oliveberg, M; Raleigh, D. P.; Wittung-Stafshede, P.; Radford, S. E.; Jackson, S. E.; Sosnick, T. R.; Marqusee, S.; Davidson, A. R.; Plaxco, K. W. Protein Folding: Defining a "Standard" Set of Experimental Conditions and a Preliminary Kinetic Data Set of Two-state Proteins. *ProteinSci*.**2005**, *14*, 602-616.
4. Prabhu, N. P.; Bhuyan, A. K. Prediction of Folding Rates of Small Proteins: Empirical Relatins Based on Length, Secondary Structure Content, Residue Type and Stability. *Biochemistry***2006**, *45*, 3805-3812.
5. Istomin, A. Y.; Jacobs, D. J.; Livesay, D. R. On the Role of Structural Class of a Protein with Two-state Folding Kinetics in Determining Correlations Between its Size, Topology, and Folding Rate. *ProteinSci*. **2007**, *16*, 3564-2569.
6. De Sancho, D.; Doshi, U.; Munõz, V. Protein Folding Rates and Stability: How Much is there Beyond Size. *J*. *Am*. *Chem*. *Soc*. **2009**, *131*, 2074-2075.
7. Campos, L. A.; Sadqi, M.; Liu, J.; Wang, X.; English, D. S.; Munõz, V. Gradual Disordering of the Native State on a Slow Two-state Folding Protein Monitored by

Single-molecule Fluorescence Spectroscopy and NMR. *J. Phys. Chem. B117*, 13120-13131.

8. Naganathan, A. N.; Perez-Jimenes, R.; Sanchez-Ruiz, J. M.; Munõz, V. Robustness of Downhill Folding: Guidelines for the Analysis of Equilibrium Folding Experiments on Small Proteins. *Biochemistry***2005**, *44*, 7435-7449.
9. Munõz, V.; Sanchez-Ruiz, J. M. Exploring Protein Folding Ensembles: A Variable-barrier model for the Analysis of Equilibrium Unfolding Experiments. *Proc. Natl. Acad. Sci. USA***2004**, *101*, 17646-17651.
10. Naganathan, A. N.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Munõz, V. Estimation of Protein Folding Free Energy Barriers from Calorimetric Data by Multi-model Bayesian Analysis. *Phys. Chem. Chem. Phys.***2011**, *13*, 17064-17076.
11. Oliva, F. Y.; Munõz, V. A Simple Thermodynamic Test to Discriminate Between Two-state and Downhill Folding. *J. Am. Chem. Soc.* **2004**, *126*, 8596-8597.
12. Ikai, A.; Fish, W. W.; Tanford, C. Kinetics fo Unfolding and Refolding of Proteins. II. Results for cytochrome *c*. *J. Mol. Biol*. **1973**, *73*, 165-184.
13. Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Munõz, V. Experimental Identification of Downhill Protein Folding. *Science***2002**, *298*, 2191-2195.
14. Fung, A.; Li, P; Godoy-Ruiz, R.; Sabchez-Ruiz, J. M.; Munõz, V. Exploring the Realm of Ultrafast Protein folding: gpW, a Midsize Natural Single-Domain with α+β Topology that Folds Downhill. *J. Am. Chem. Soc.* **2008**, *130*, 7489-7495.
15. Naganathan, A. N.; Li, P.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Munõz, V. Navigating the Downhill Protein Folding Regime via Structural Homologues. *J. Am. Chem. Soc*. **2010**, *132*, 11183-11190.

16. Greene, R. F.; Pace, C. N. Urea and Guanidine Hydrochloride Denaturation of Ribonuclease, Lysozyme, Alpha-chymotrypsin, and Beta-lacto globuline. *J. Biol. Chem*. **1974**, *249*, 6388-5393.
17. Gupta, R.; Yadav. S.; Ahmad, F. Protein Stability: Urea-induced versus Guanidine-induced Unfolding of Metmyoglobin. *Biochemistry***1996**, *35*, 11925-11930.
18. Gupta, R.; Ahmad, F. Protein Stability: Functional Dependence of Denaturational Gibbs Energy on Urea Concentration. *Biochemistry***1999**, *38*, 2471-2479.
19. Ferreon, A. C. M.; Bolen, D. W. Thermodynamics of Denaturant-induced Unfolding of a Protein that Exhibits Variable Two-state Denaturation. *Biochemistry***2004**, *43*, 13357-13369.
20. Pace, C. N.; Vanderburg, K. E. Determining Globular Protein Stability: Guanidine Hydrochloride Denaturation of Myoglobin. *Biochemistry***1979**, *18*, 288-292.
21. Santoro, M. M.; Bolen, D. W. A Test of the Linear Extrapolation of Unfolding Free Energy Changes over an Extended Denaturant Concentration Range. *Biochemistry***1992**, *31*, 4901-4907.
22. Bai, Y.; Sosnick, T. R.; Mayne, L.; Englander, S. W. Protein Folding Intermediates: Native-state Hydrogen Exchange. *Science***1995**, *269*, 192-197.
23. Englander S. W. Protein Folding Intermediates and Pathways Studied by Hydrogen Exchange. *Annu. Rev. Biophys. Biomol. Struct*.**2000**, *29*, 213-238.
24. Yasin, U. M.; Sashi, P.; Bhuyan, A. K. Expansion and Internal Friction in Unfolded Protein Chain. *J. Phys. Chem. B***2013**, *117*, 12059-12064.
25. Ibarra-Molero, B.; Sanchez-Ruiz, J. M. Are There Equilibrium Intermediate States in the Urea-induced Unfolding of Hen Egg White Lysozyme? *Biochemistry***1997**,

36, 9616-9624.

26. Bieri, O.; Wildegger, G.; Bachmann, A.; Wagner, C.; Kiefhaber, T. A Salt-induced Kinetic Intermediate is on a New Parallel Pathway of Lysozyme Folding. *Biochemistry***1999**, *38*, 12460-12470.
27. Chen, L.; Hodgson, K. O.; Doniach, S. A Lysozyme Folding Intermediate Revealed by Solution X-ray Scattering. *J. Mol. Biol*. **1996**, *261*, 658-671.
28. Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins. *Science***1991**, *254*, 1598-1603.
29. Naganathan, A. N.; Doshi, U.; Munõz, V. Protein Folding Kinetics: Barrier Effects in Chemical and Thermal Denaturation Experiments. *J. Am. Chem. Soc*. **2007**, *129*, 5673-5682.
30. Yang, Y.; Hamaguchi, K. Hydrolysis of 4-methylumbelliferyl N-acetyl-chitotroside Catalyzed by Hen and Turkey Lysozymes. *J. Biochem*. **1980**, *87*, 1003-1014.
31. Altieri, A. S.; Hinton, D. P.; Byrd, R. A. Association of Biomolecular Systems *via* Pulsed Field Gradient NMR Self-diffusion Measurements. *J. Am. Chem. Soc*. **1995**, *117*, 7566-7567.
32. Jones, J. A.; Wilkins, D. K.; Smith, L. J.; Dobson, C. M. Characterization of Protein Unfolding by NMR Diffusion Measurements. *J. Biomol. NMR***1997**, *10*, 199-203.
33. Urry, D. W. Protein-heme Interactions in Heme-proteins: Cytochrome *c*. *Proc. Natl. Acad. Sci. USA***1965**, *54*, 640-643.
34. Sadqi, M; Fushman, D.; Munõz, V. Atom-by-atom Analysis of Global Downhill Protein Folding. *Nature***2006**, 442, 317-321.

35. Rao, M. T.; Bhuyan, A. K.; Venu, K.; Sastry, V. S. S. Nonlinear Effect of GdnHCl on Hydration Dynamics of Proteins: A ^{1}H Magnetic Relaxation Dispersion Study. *J. Phys. Chem. B***2009**, *113*, 6994-7002.
36. Santoro, M. M.; Bolen, D. W. Unfolding Free Energy Changes Determined by the Linear Extrapolation Method. 1. Unfolding of Phenylmethylsulfonyl α-chymotrypsin Using Different Denaturants. *Biochemistry***1988**, *27*, 8063-8068.
37. Tanford, C. Protein Denaturation. *Adv. Protein. Chem.* **1970**, *24*, 1-95.
38. Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely Slow Intramolecular Diffusion in Unfolded Protein L. *Proc. Natl. Acad. Sci. USA***2010**, *107*, 13713-13717.
39. Voelz, V. A; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. Unfolded-State Dynamics and Structure of Protein L Characterized by Simulation and Experiment. *J. Am. Chem. Soc.* **2010**, *132*, 4702-4709.
40. Lapidus, L. J. (2013) Exploring the Top of the Protein folding Funnel by Experiments. *Curr. Opin. Struct. Biol.* **2013**, *23*, 30-35.
41. Kitao, A.; Hayward, S.; Go, N. Energy Landscape of a Native Protein: Jumping-among-minima Model. *PROTEINS***1998**, *33*, 496-517.
42. Whitten, S. T.; Garcia-Monero, E. B.; Hilser, V. J. Local Conformational Fluctuations can Modulate the Coupling between Proton Binding and Global Structural Transitions in Proteins. *Proc. Natl. Acad. Sci. USA***2005**, *102*, 4282-4287.
43. Tavernelli, I.; Cotesta, S.; Di Iorio, E. E.Protein Dynamics, Thermal Stability, and Free-energy Landscapes: A Molecular Dynamics Investigation. *Biophys. J.* **2012**, *85*, 2641-2649.

44. Cotesta, S.; Tavernelli, I.; Di Iorio, E. E. Dynamics of RNase-A and S-protein: A Molecular Dynamics Simulation of the Transition toward a Folding Intermediate. *Biophys*. *J*. **2012**, *85*, 2633-2640.
45. Krishna, M. M.; Hoang, L.; Lin, Y.; Englander, S. W. Hydrogen Exchange Methods to Study Protein Folding. *Methods***2004**, *34*, 51-64.
46. Bai, Y. Protein Folding Pathways Studied by Pulsed- and Native-state Hydrogen Exchange. *Chem*. *Rev*. **2006**, *106*, 1757-1768.
47. Bernstein, R.; Schmidt, K. L.; Harbury, P. B.; Marqusee, S. Structural and Kinetic Mapping of Side-chain Exposure onto the Protein Energy Landscape. *Proc*. *Natl*. *Acad*. *Sci*. *USA***2011**, *108*, 10532-10537.
48. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) The Folding of Hen Lysozyme Involves Partially Structured Intermediates and Multiple Pathways. *Nature***1992**, *358*, 302-307.
49. Miranker, A.; Robinson, C. V.; Radford, S. E.; Aplin, R. T.; Dobson, C. M. Detection of Transient Protein Folding Populations by Mass Spectrometry. *Science***1993**, *262*, 896-900.
50. Kulkarni, S. K.; Ashcroft, A. E.; Carey, M.; Masselos, D.; Robinson, C. V.; Radford, S. E. A Near-native State on the Slow Refolding Pathway of Hen Lysozyme. *Protein*. *Sci*. **1999**, *8*, 35-44.
51. Korzhnev, D. M.; Kay, L. E. Probing Invisible, Low-populated States of Protein Molecules by Relaxation Dispersion NMR Spectroscopy: an Application to Protein Folding. *Acc*. *Chem*. *Res*. **2008**, *41*, 442-451.
52. Frauenfelder, H.; Young, R. D.; Fenimore, P. W. Dynamics and the Free-energy Landscape of Proteins, Explored with the Mössbauer Effect and Quasi-elastic

Neutron Scattering. *J. Phys. Chem. B***2013**, *117*, 13301-13307.

53. Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the Free Energy Surface for Protein Folding with Single Molecule Fluorescence Spectroscopy. *Nature***2002**, *419*, 743-747.
54. Sherman, E.; Haran, G. Coil-Globule Transition in the Denatured State of a Small Protein. *Proc. Natl. Acad. Sci. USA***2006**, *103*, 11539-11543.
55. Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.;, Horsley, D. A.; Seckler, R.; Bakajin, O. *etal*. Mapping Protein Collapse with Single-Molecule Fluorescence and Kinetic Synchrotron Radiation Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 105-110.
56. Huang, F.; Sato, S.; Sharpe, T. D.; Ying, L.; Fersht, A. R. Distinguishing between Cooperative and Unimodal Downhill Protein Folding. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 123-127.
57. Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. Characterizing the Unfolded States of Proteins Using Single-Molecule FRET Spectroscopy and Molecular Simulations. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 1528-1533.
58. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 2649-2654.
59. Ziv, G.; Haran, G. Protein Folding, Protein Collapse, and Tanford's Transfer Model: Lessons From Single Molecule FRET. *J. Am. Chem. Soc.* **2009**, *131*, 2942-2947.
60. Ferreon, A. C. M.; Deniz, A. A. Protein Folding at Single-Molecule Resolution.

*Biochim. Biophys. Acta***2011**, *1814*, 1021-1029.

61. Haran, G. How, When and Why Proteins Collapse: The Relation to folding. *Curr. Opin. Struct. Biol.* **2012**, *22*, 14-20.

2.7 **Supporting Information**

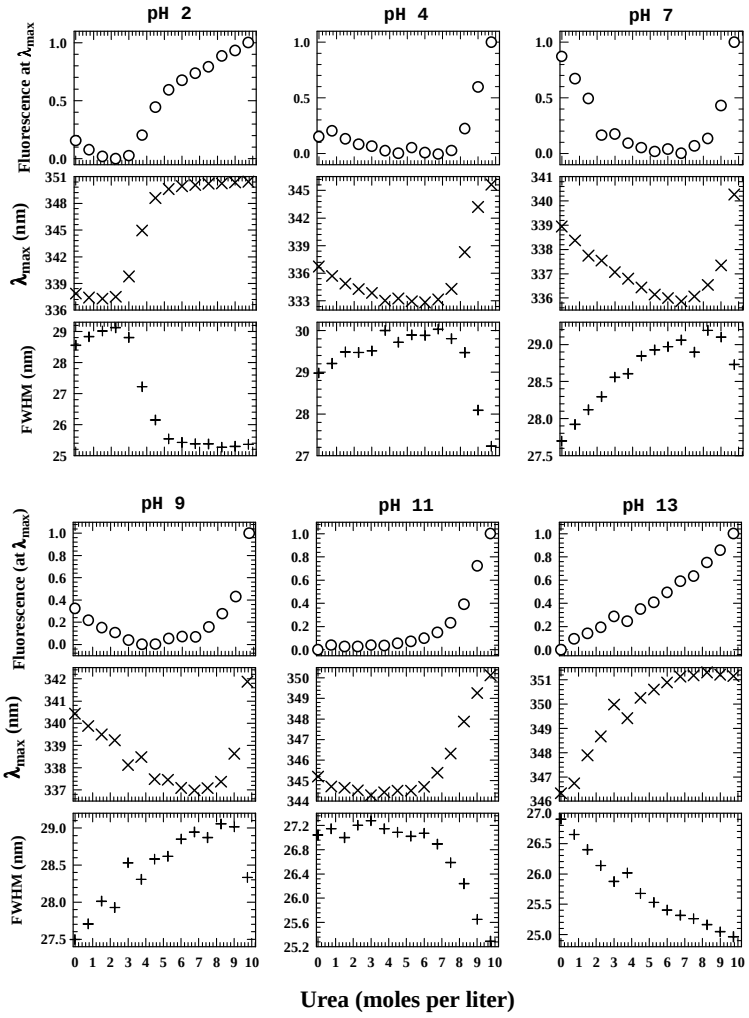


Fig. S1. Urea dependence of tryptophan fluorescence intensity, emission wavelength maximum, and the full width at half maximum (FWHM) of the emission spectrum at indicated values of pH. Incomplete unfolding at pH 4, 7, 9, and 11 within the solubility limit of urea is evident.

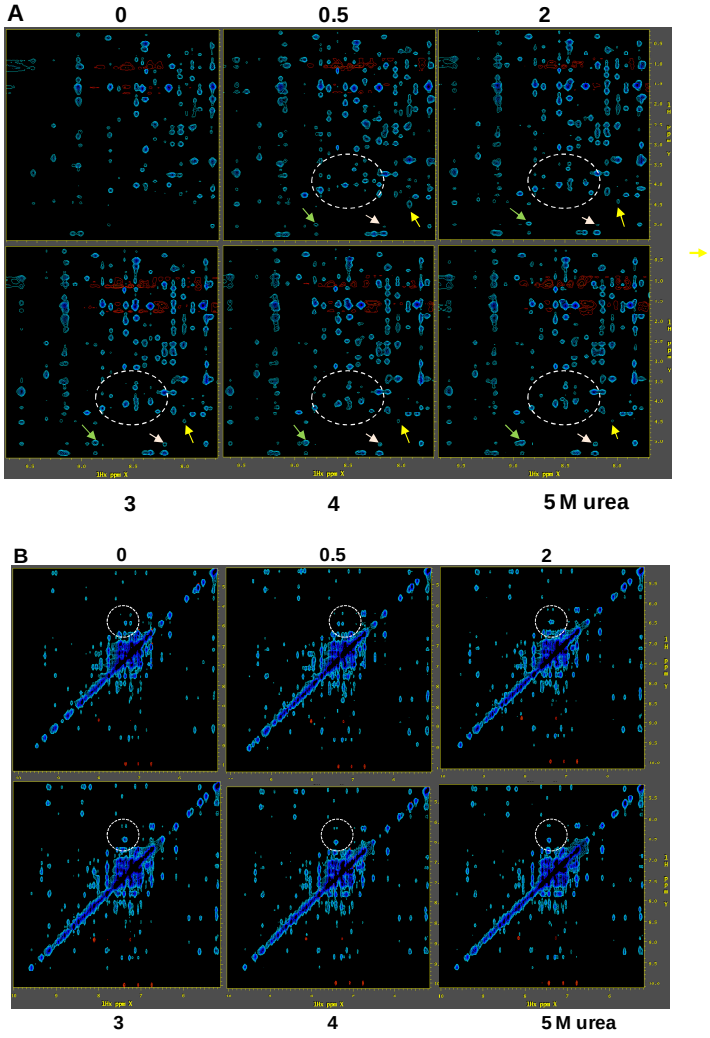


Fig. S2. Screen shots of regions of NOESY spectra of lysozyme at indicated urea concentrations, pH 5, and 25°C. The crowded aliphatic region is not shown. Some of the minor changes at the indicated subdenaturing concentrations of urea are shown by arrowheads and circles.

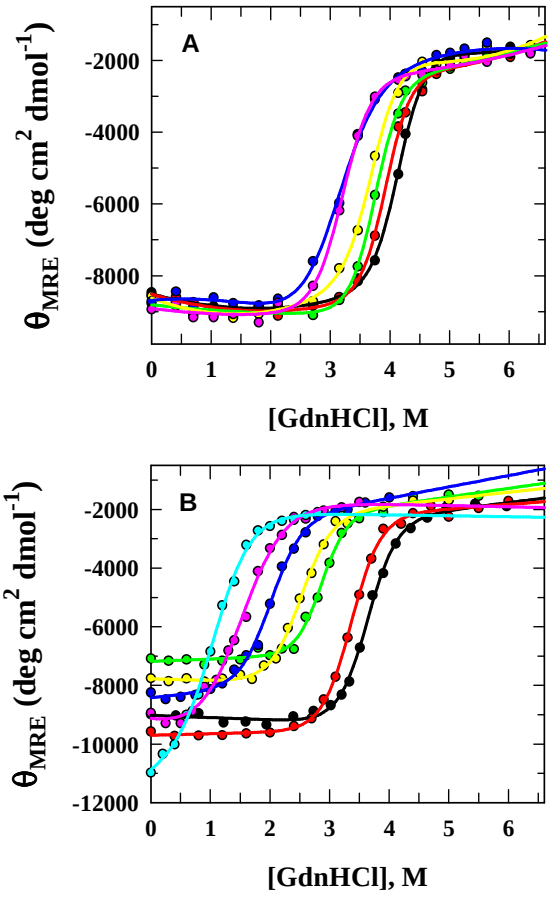


Fig. S3. Far-UV monitored GdnHCl-induced equilibrium unfolding of lysozyme at different concentrations of urea, 20 mM sodim acetate, pH 5, 25°C. (*A*) 0 M (●), 0.55 M (●), 1.1 M (●), 2.3 M (●), 2.7 M (●), and 2.8 M (●) urea. (*B*) 1 M (●), 2.1 M (●), 3.2 M (●), 4.2 M (●), 5.0 M (●), 5.1 M (●), and 7.3 M (●) urea. The fits through data are according to the two state (N⇌U) transition assuming linear dependence of Δ*G* on GdnHCl molarity.

Chapter-3

Disorder-to-Order Collapse of the Intrinsically Disordered Protein *At*PP1 Under Subdenaturing Conditions of Urea: A ^{15}N NMR Relaxation Study

3.1 Abstract

Conformationally ill-defined regions of most of intrinsically disordered proteins (IDPs) are known to undergo disorder-to-order transition upon binding to other proteins and ligands. Many IDPs also exhibit the propensity to undergo chain collapse transitions in aqueous milieu. This study uses ^{15}N NMR relaxation-based backbone dynamics to demonstrate collapse and disorder-to-order transition for the IDP *Arabidopsis thaliana* phloem protein 1 (*At*PP1) when it is placed under subdenaturing solution conditions created by low levels of urea. Nonspecific binding of urea to the protein dramatically reduces the backbone fluctuations giving rise to rigidity. The rotational correlation time of *At*PP1 marginally decreases from 9.1(±0.4) ns in the absence of urea to 7.9(±0.7) in the presence of 0.4 M urea. Polyfunctional interactions of urea molecules to protein atoms from different parts may serve to produce intramolecular cross-links forcing the protein to collapse. Dynamic stiffening and dimensional shrinkage in the presence of subdenaturing levels of urea should be general effects for all proteins, but are more pronounced in the case of IDP because of the high measurable contrast in dynamic disorder between the initial natively folded state and the urea induced collapsed state.

3.2 Introduction

Disorder in some protein structures and their importance in functional regulation have been described since the early days of X-ray crystallography.[1] But disorder in solution structure of many proteins has been most decisively established in later years by NMR spectroscopy which is more committed to describing conformational and dynamic disorder.[2-13] Proteins containing disordered regions, generically termed intrinsically disordered proteins (IDPs), are widespread in the cellular milieu accounting for at least a third of human proteins.[14] The disorder is believed to be required for binding with interacting proteins and ligands. A large body of reports have provided evidence that binding with interacting ligands leads to folding of the disordered regions,[3-5,15-20] although some studies have found significant disorder in IDPs even when ligands are bound to them.[21-23] In addition to serving as effector molecules for the interacting ligands, IDPs have diverse functional roles, for all of which the disordered region required in some way or the other.[24]

An important aspect regarding folding and solution properties of IDPs is their propensity to collapse in water.[25-32] The net charge per residue appears to be a major factor for the tendency to collapse,[29] which has been observed when IDP charges are screened by solution ions dissociated from simple salt or non-denaturing levels of the ionic denaturant GdnHCl.[33] IDP collapse has also been reported in the presence of the ionic detergent SDS.[30] In view of the myriad of structural, folding, and functional importance of IDP, a detailed understanding of the processes involving their folding upon binding, disorder-to-order transition, and chain collapse is imperative.

This study describes disorder-to-order and collapse transitions of the IDP *At*PP1, a phloem protein from *Arabidopsis thaliana*. Phloem proteins are suspected to

be involved in a number of processes, including the transport and delivery of nutrients, hormones, and RNA to the phloem translocation system.[34] In fact, no structural and functional characterization of *At*PP1 has been reported till date, and the putative functions listed are based on studies on the homologous *Cucurbita maxima* phloem protein 16 (*Cm*PP16). It was therefore sought to determine the solution structure of *At*pp1, in the course of which the protein has been found to be an IDP. Deferring the structural description here, the amide resonance assignments are used for backbone relaxation studies using T_1, T_2, and $\{^1H\}$-^{15}N NOE. Relaxation parameters indicate that when *At*PP1 is subdentured in the presence of very low urea levels, it collapses and undergoes a disorder-to-order transition. It is argued that all proteins must stiffen and shrink due to nonspecific binding interaction with urea when the latter is used at subdenatuirng concentrations. Dynamic and dimensional changes associated with IDP collapse are enhanced and unmistakable because unlike folded single-domain native proteins, the initial natively folded state of IDP is largely disordered.

3.3 Experimental Section

Procedures for overexpression and ^{15}N labeling of *At*pp1 have been described in the next chapter. NMR samples were ~150 µM in protein concentration prepared in 7 mM sodium acetate buffer, 10% D_2O, pH 3.5. All experiments were performed at 25°C in a 500 MHz Bruker spectrometer using a 5 mm triple resonance inverse probe equipped with actively shielded field gradient coils.

T_1, T_2, and NOE experiments were performed by sequences appropriate for inversion recovery, CPMG, and steady-state $\{^1H\}$–^{15}N NOE,[35] respectively. Eight

inversion recovery delays from 10 ms to 1200 ms were used for T_1 measurements. For T_2 measurements 7 CPMG delays with multiple of 16.96 ms from 16.96 ms to 169.6 ms were employed. Spectra were recorded as 128×2048 complex matrices with 32 scans per complex t1 point. Steady-state NOE spectra with and without proton saturation were recorded as 256×1024 complex matrices with 64 scans for each complex t1 point. Proton presaturation in NOE experiments was achieved by the use of the standard 120° ^{1}H pulses.[36] In all experiments echo-antiecho was used in the indirect dimension. To extract the relaxation time constants T_i, peak intensities of all resolved resonances in T_1 and T_2 spectra were plotted with relevant recovery delays, and fitted to single exponential decays

$$I(t) = Ae^{-\frac{t}{T_i}} \quad (1)$$

Values of $\{^1H\}$–^{15}N NOE were determined from

$$\mathrm{NOE} = \frac{I_{\mathrm{sat}}}{I_{\mathrm{unsat}}} \quad (2)$$

where, I_{sat} and I_{unsat} are resonance intensities with and without proton saturation, respectively.

3.4 Results

3.4.1 ANS Binding, Structural Collapse and Equilibrium Unfolding. Disorder in the structure of *At*PP1 has been evident to us because of the difficulty of resolving a large number of side-chain and a few main-chain resonances in multidimensional NMR experiments that have been performed to determine the solution structure of the protein. For the present, the result for the basic experiment of binding of the dye ANS (8-anilino naphthalene sulfonate) to the native protein is

shown in Figure 1a. The idea is ANS binding would be substantial if the protein is structurally disordered allowing larger accessibility of the dye toward hydrophobic surfaces in the protein interior.[37,38] As the protein is titrated with urea up to 1 M in the presence of constant 5 μM ANS, the ANS fluorescence decreases to half with a concomitant red shift, (Figure 1a), suggesting a urea-mediated transition by which the protein acquires structure and turns compact. This transition is reminiscent of a disorder-to-order transition, and provides a basis to think that *At*PP1 is an IDP.

To examine protein stability, urea denaturation was monitored by intrinsic tryptophan fluorescence and 228-nm CD absorption in 50 mM acetate buffer, pH 3.5. The tryptophan fluorescence initially decreases dramatically in the 0-2 M urea range, and increases with further addition of the denaturant (Figure 1b). The sharp initial decrease has no baseline, and is clearly atypical of a cooperative transition (Figure 1c). The fluorescence increase at higher urea (>2 M) appears to be due to structure unfolding, because the CD-monitored unfolding transition, which is reversible and cooperative and has clear pre- and post-transition baselines, provides a transition midpoint of ~2.74 M urea ($\Delta G^{\circ}\sim$ 7.5(±0.1) kcal mol^{-1}). Hence, the intrinsic protein fluorescence decrease in the 0-1 M range of urea (inset, Figure 1c) together with ANS fluorescence quenching in the same range of the denaturant can be taken as primary indications of structural collapse of intrinsically disordered *At*PP1.

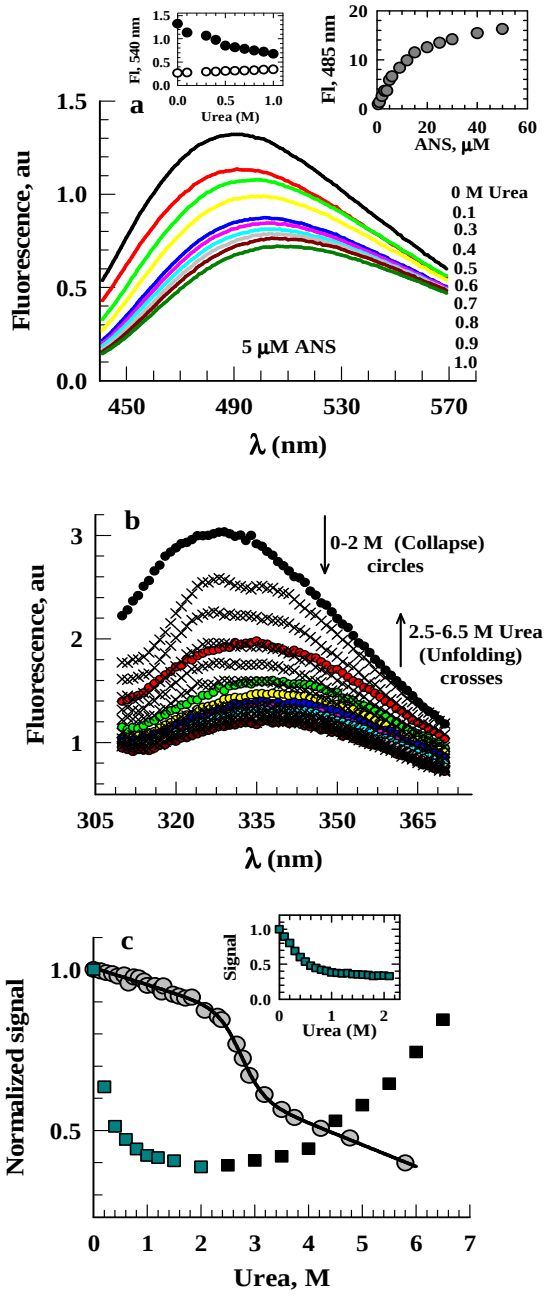


Figure 1. Basic results for IDP attribute and urea-mediated collapse of *At*PP1. (a) ANS binding to *At*PP1 in the 0-1 M range of urea. Red-shift and fluorescence decrease with higher urea in this range indicates a structural collapse. The left *inset* shows urea-dependent decreases in ANS fluorescence with protein (solid symbols). The data in the absence of the protein (open symbols) serves as the control. The right *inset* depicts ANS titration of the native protein, i.e., in the absence of urea. (b) Tryptophan emission spectra are red-shifted with decreasing fluorescence as urea level is incremented from 0 to 2 M (solid colored symbols), but are blue-shifted with increasing fluorescence (cross symbols) at urea higher than 2 M. The fluorescence decrease and increase processes are due to initial collapse and eventual protein unfolding, respectively. (c) In urea unfolding, the sharp fluorescence decrease without a native-state baseline (cyan squares) indicates a protein collapse. This phase is also shown in the *inset*. At higher urea, the fluorescence increase (black squares) suggests structure unfolding, albeit a clear unfolded baseline is not obtained. The far-UV CD-monitored urea titration (grey circles) provides the evidence for a cooperative unfolding process. The solid line represents a two-state fit of data with transition midpoint C_m ~2.74 M urea, and $\Delta G°$~7.5 kcal mol^{-1}.

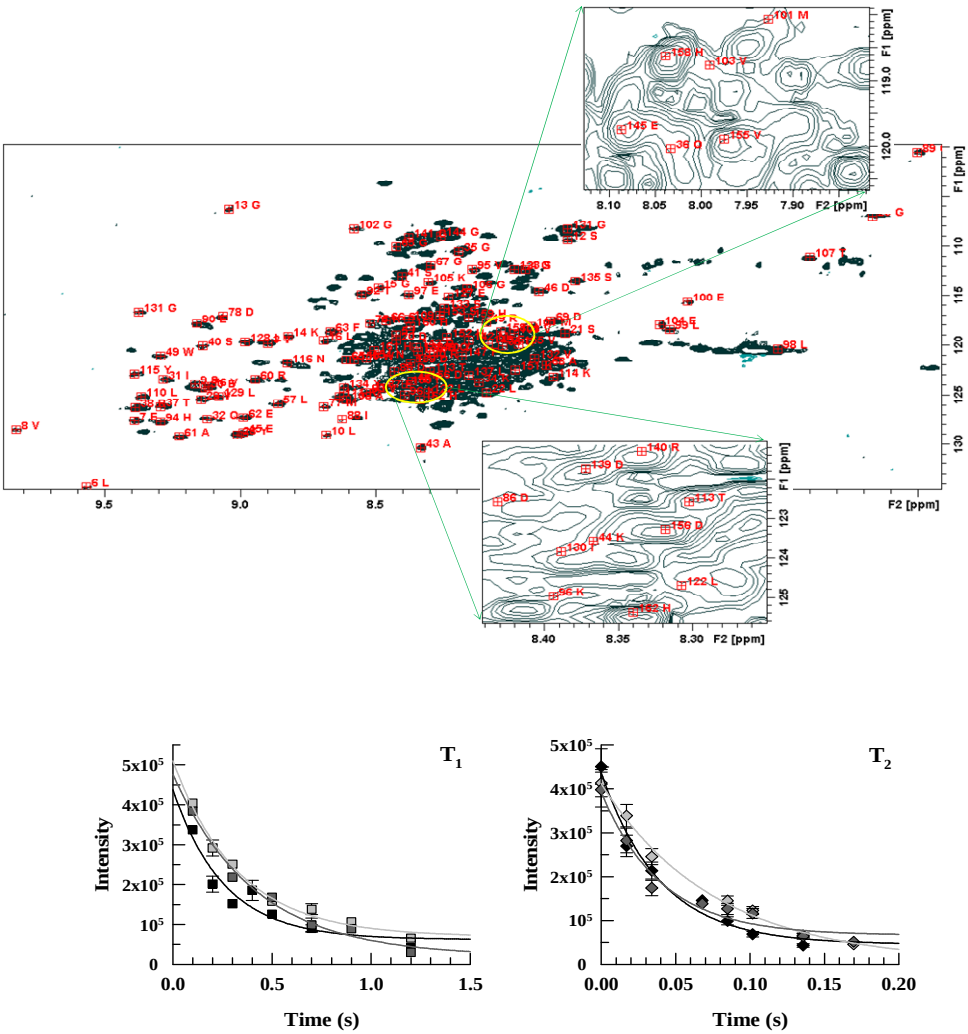


Figure 2. A representative HSQC spectrum used for determination of T2 (84.8 ms CPMG delay). Curves for arbitrarily chosen resonances are also shown. Solid lines are least squares fits to exponential function for extracting (a) T_1 and (b) T_2.

3.4.2 T_1 and T_2 of Backbone ^{15}N. Standard pulse schemes[35,39] were used to record ^{1}H-^{15}N correlation spectra, from which T_1 and T_2 time constants are readily extracted by exponential fit of the decaying cross peak intensity with incremented relaxation delay (Figure 2). Although one expects 151 backbone amide cross peaks for the sequence of 156 residues with 4 prolines, some may not appear because of exchange broadening. Even amongst the observable cross peaks in the ^{1}H-^{15}N spectrum, only the non-overlapping ones could be considered useful. Thus, intensities of only 93, 70, and 54 cross peaks for spectra corresponding to 0, 0.2, and 0.4 M urea, respectively, were quantified for further analysis. A survey of T_1 values across the sequence (Figure 3) does not provide any direct information about backbone dynamics, but the values in general appear to decrease a little that at best may suggest some increase in high-frequency motions. The T_2 values for C-terminal residues are clearly larger irrespective of the content of urea (Figure 3). Because smaller values of transverse relaxation times are related to chemical or conformational exchange processes whose frequencies are higher than the CPMG repetition rates,[39]large amplitude exchange motions are not prevalent in the C-terminal segment of about 25 residues. This less flexible stretch singles out from the rest of the sequence, and its dynamic rigidity relative to rest of the residues in the sequence is preserved even when the protein is subdenatured in the presence of 0.2 or o.4 M urea (Figure 3).

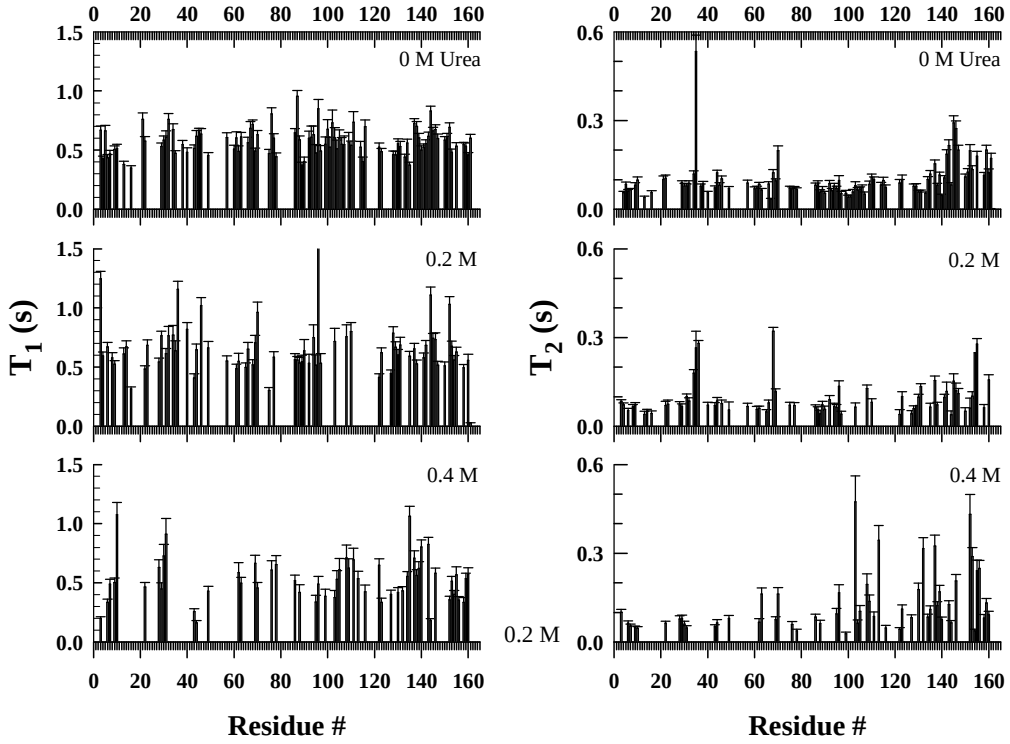


Figure 3. T1 and T2 values of ^{15}N resonances for resolved residues across the sequence at indicated urea concentrations.

The T_1/T_2 ratio can be used to estimate the global correlation time (τ_c) of the protein subject to conditions that fast internal motions affect both T_1 and T_2 to the same extent and slow conformational averaging processes do not shorten the T_2 values. Adopting the statistical criteria for selection of residues whose T_1/T_2 ratio is useful for τ_c estimation,[40,41] the overall correlation time of *At*PP1 was estimated by

$$\tau_c = \frac{1}{2\omega_N}\sqrt{\frac{6T_1}{T_2} - 7} \quad (3)$$

where ω_N is the Larmor frequency of ^{15}N (=50.68 MHz). The τ_c value of 9.1 ns for the native state decreases to 8.4 and 7.9 ns in the presence of 0.2 and 0.4 M urea, respectively (Figure 4), suggesting that the overall dimension of the protein shrinks under increasingly subdenaturing conditions.

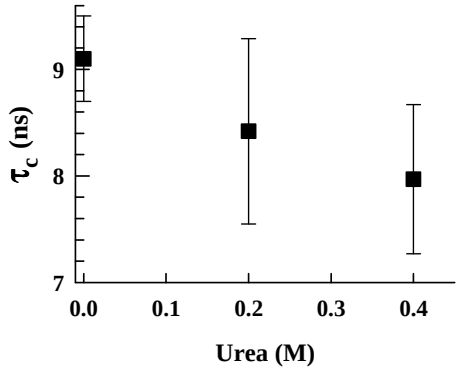


Figure 4. Decrease in the global correlation time τ_c as the protein is subdenatured/collapsed in the presence of 0.2 and 0.4 M urea.

3.4.3 {^{1}H}-^{15}N NOE. The measured heteronuclear NOE for residues across the sequence of *At*PP1 occasionally shows a few values exceeding 0.82 (Figure 4). Considering contributions of both ^{1}H-^{15}N dipolar interactions and chemical shift anisotropy, the NOE value at ω_N=50.68 MHz varies from -3.6 under extreme narrowing conditions ($\omega_N\tau_C << 1$) to 0.82 when τ_C well-exceeds the Larmor frequency,[39,42] meaning excess NOE values for a few residues here is due to measurement error (Figure 5). Residues that register smaller NOE values are generally highly mobile, where the internal correlation time also contributes sizably. Absence of negative NOEs indicates no overwhelming contribution of rapid internal motions at any measured site. The pattern of variation of NOEs across the protein backbone is not strictly maintained amongst different subdenatured states of the protein. For example, NOE values for the last 20 C-terminal residues are uniformly

smaller for the protein in water when compared with those in 0.2 and 0.4 M urea solutions (Figure 5), which qualitatively suggest stiffening of these C-terminal residues in the subdenatured state of the protein. This inference is consistent with T_1 and T_2 results obtained above that subdenatured states produced with increments of the denaturant are dynamically constrained.

3.4.4 Model-Free Analysis of Relaxation Data. Procedures for extraction of parameters defining the motion of the $^{1}H–^{15}N$ bond vector and internal and global correlation times are well developed.[39,43,44,45,46,47,48,49] Briefly, the strengths of the two relaxation mechanisms for the backbone ^{15}N spin, namely, dipolar coupling to the directly bonded amide proton and chemical shift anisotropy, are incorporated in the expressions relating spectral density functions and T_1, T_2, and NOE enhancement by[39,42]

$$\frac{1}{T_1} = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \quad (4)$$

$$\frac{1}{T_2} = \frac{d^2}{2}[4J(0)J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{6}[3J(\omega_N) + 4J(0)] \quad (5)$$

$$\mathrm{NOE} = 1 + \left[\left(\frac{\gamma_H}{\gamma_N}\right) d^2\{6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\}T_1\right] \quad (6)$$

where $d = \frac{\mu_0}{4\pi}\gamma_H\gamma_N\frac{h}{2\pi}\langle\frac{1}{r_{NH}{}^3}\rangle$ and $c = \frac{\omega_N}{\sqrt{3}}(\sigma_\parallel - \sigma_\perp)$, with μ_o, and γ_H and γ_N being the permeability of free space, and gyromagnetic ratios of ^{1}H and ^{15}N, respectively. The N-H bond length, r_{NH}≈1.02 Å. The difference of the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor as applicable to the

peptide bond is generally taken as -160 ppm.[39,50] When the overall rotational diffusion is isotropic ($D_{\parallel}=D_{\perp}=D$), the relaxation data are conveniently analyzed by expressing the spectral densities at angular frequencies ω_i, $J(\omega_i)$, in terms of minimal motional parameters according to Lipari and Szabo[43,44]

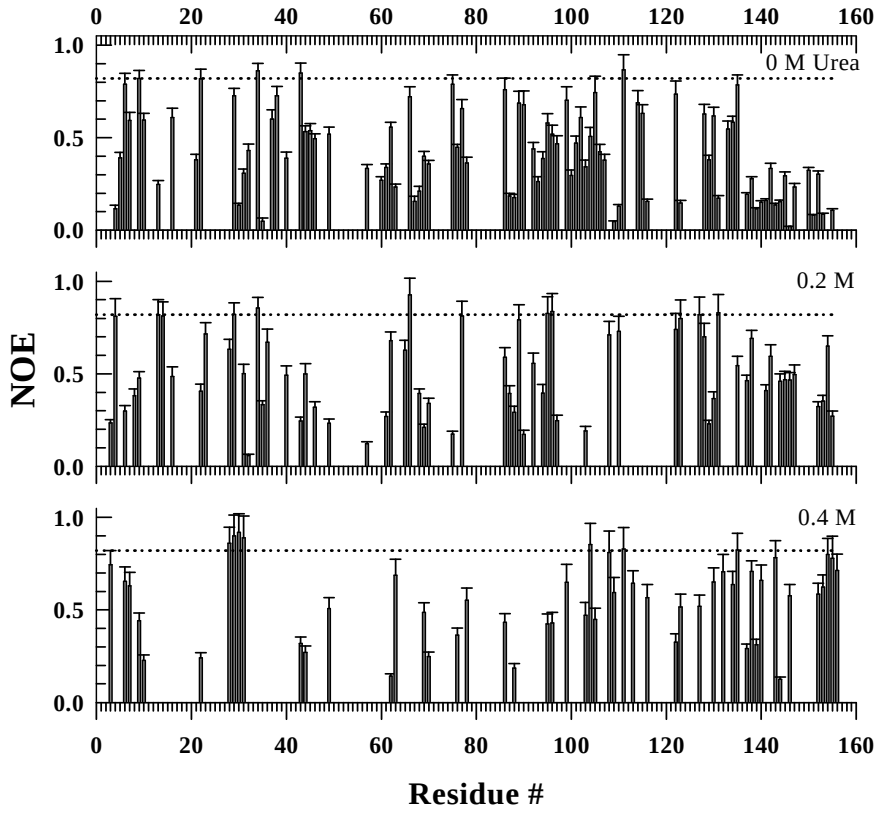


Figure 5. $\{^1H\}-^{15}N$ heteronuclear NOE for residues resolved in the presence of 0, 0.2, and 0.4 M urea. The dotted line in each panel indicates the maximum value of NOE possible, which is 0.82. Any value larger than 0.82 is due to measurement error.

$$J(\omega_i)=\left\{\frac{S^2\tau_c}{1+\omega_i^2\tau_c^2}\right\}+\left\{\frac{(1-S^2)\tau}{1+\omega_i^2\tau^2}\right\} \quad (7)$$

where S is the generalized order parameter describing the amplitude of internal motion of the N-H bond, $1/\tau= 1/\tau_c +1/\tau_e$ with τ_e representing the correlation time for rapid internal motion. A chemical exchange or conformational exchange parameter, $1/T_{ex}$ ($=R_{ex}$) is sometimes added to improve the overall agreement between theory and experiments. Thus, the $J(\omega_i)$ function can have a minimum of two (S^2 and τ_c) and a maximum of four (S^2, τ_c, τ_e, and $1/T_{ex}$) dynamical parameters. Which of the four $J(\omega_i)$ functions best describe the relaxation parameters for each residue of the protein is determined by using the Model-free program developed by Palmer and coworkers.[47] This analysis is convenient for an axially symmetric protein undergoing isotropic rotational diffusion with defined components of the diffusion tensor parallel and perpendicular to the principal axis of the symmetry.

The S^2 values across the *At*PP1 sequence extracted from relaxation data using the Model-free program are plotted in Figures 6. Strikingly, for the protein in aqueous medium all residues across the sequence, except R38, L122, and H162 from the His_6 tag, register unusually low values, actually approaching zero for some, suggesting a highly mobile backbone. Although the result is not unexpected for an IDP like *At*PP1, unusually low values of S^2 is remarkable. Even more interesting is the dramatic increase of S^2 values under increasingly subdenaturing conditions created by adding urea. Average values of S^2 are 0.12, 0.59, and 0.82 for 0, 0.2, and 0.4 M urea, respectively, suggesting gradual dynamic constraints on the protein chain. The high degree of reduced mobility, essentially a disordered-to-ordered structural transition, appears to suggest denaturant-induced collapse of the *At*PP1 chain.

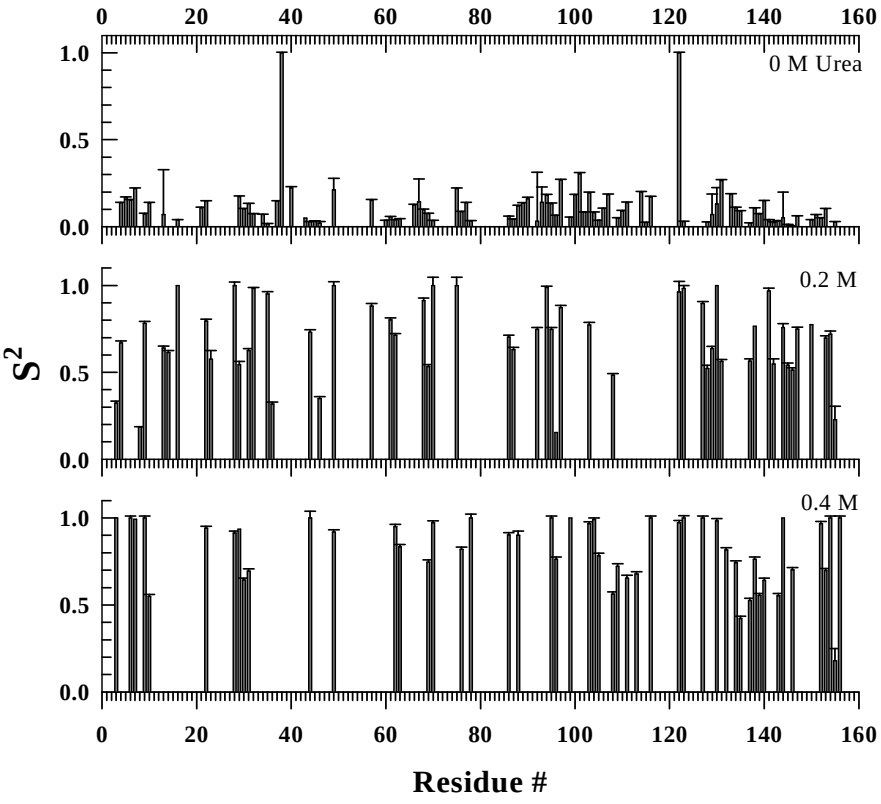


Figure 6. The square of the generalized order parameter, S^2, for resolved resonances at the indicated concentrations of urea. Values of S^2 were determined by using the Model-free program kindly provided by Arthur Palmer.

The parameter for internal correlation time, τ_e, was explicitly included to fit the relaxation data for a few residues (Figure 7a-c). The values, which fall in the 0.5-9 ps range in the absence of urea, are much larger (17 to 165 ps) in the presence of 0.4 M urea, notwithstanding the lack of residue-wise correspondence. The significance of

this result is not considered here, because the interpretation of τ_e is not straightforward especially in the absence of a well-defined motional model. The inclusion of the R_{ex} term was not needed for any residue under 0 M urea condition, but its inclusion for a few residues in the subdenatured states improved the overall fit (Figure 7d,e). Even though the accuracy and significance of these values are often not certain, residues associated with R_{ex} are generally taken to undergo chemical exchange and other pseudo-first order processes that contribute to the transverse magnetization decay.

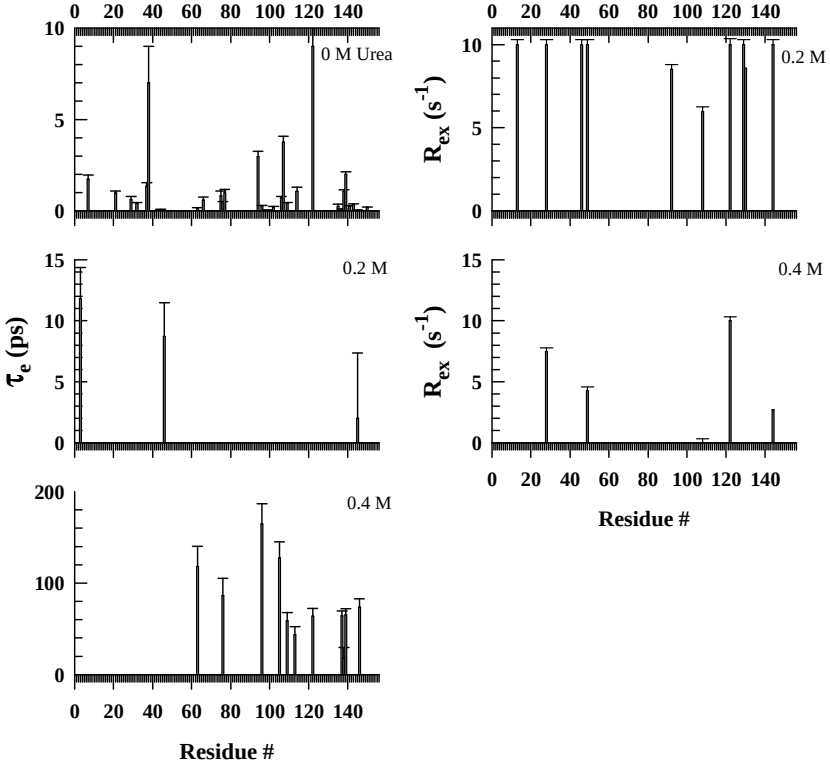


Figure 7. Internal correlation time, τ_e, and chemical exchange rate, R_{ex}, for those residues which required these parameters for analyzing the respective relaxation data. Inclusion of R_{ex} was not necessary for any residue when there was no urea in the solution.

3.4.5 Reduced Spectral Density Mapping. Relaxation data for intrinsically disordered proteins like *At*PP1 that may lack a well-defined topology can be analyzed more appropriately by spectral density mapping. Anisotropic tumbling, as might occur for partly structured, denatured or unfolded proteins, gives rise to a large number of uncorrelated motions and a clear separation of internal and global motions is often not feasible. In this sense, the assumptions of the model- free formalism are not strictly valid. Hence, the method of spectral density mapping is employed where relaxation rate constants are analyzed by first determining $J(\omega)$.[51,52] When only three ^{15}N relaxation parameters are available, reduced spectral density mapping[48,51,52,53] provides a way to extract some features of motional trends and correlation times. The assumption of $J(\omega) \approx J(\omega_H+\omega_N) \approx J(\omega_H-\omega_N)$ at high frequencies provides the following relations

$$J(0)=\frac{3}{2(3d^2+c^2)}\left[-\frac{1}{2T_1}+\frac{1}{T_2}-\frac{3}{5}R_{NOE}\right] \quad (8)$$

$$J(\omega_N)=\frac{1}{3d^2+c^2}\left[\frac{1}{T_1}-\frac{7}{5}R_{NOE}\right] \quad (9)$$

$$J(\omega_H)=\frac{1}{5d^2}R_{NOE} \quad (10)$$

where $R_{NOE}=(\theta-1)\frac{1}{T_1}\left(\frac{\gamma_N}{\gamma_H}\right)$ with θ for the heteronuclear NOE, $c^2 \sim 0.9\times10^9$ $(rad/s)^2$ at 500 MHz of proton frequency,[54] and $d^2 \sim 1.35\times10^9$ $(rad/s)^2$.

Calculated spectral densities J(0), $J(\omega_N)$, and $J(\omega_H)$ for many residues of *At*PP1 in the presence of 0, 0.2, and 0.4 M urea are shown in Figure 8. The J(0) value is fairly high for some residues, often approaching 7 or more, which is due tocontributions from exchange motions and other pseudo-order processes to T_2. But the C-terminal residues register relatively smaller J(0) values consistently irrespective

ofthe presence of urea. Since J(0) is the dominant contributor and is inversely proportional values (Figure 3) do not appear to sample a large number of conformations and are to transverse relaxation time,[48] these residues that are already seen to have larger T_2values (Figure 3) do not appear to sample a large number of conformations and are hence relatively rigid. Slopes and intercepts of plots of J(0) and J(ω_N), and J(0) and J(ω_H) (Figure 9) can be used to estimate the correlation times, τ_i, according to[48,55]

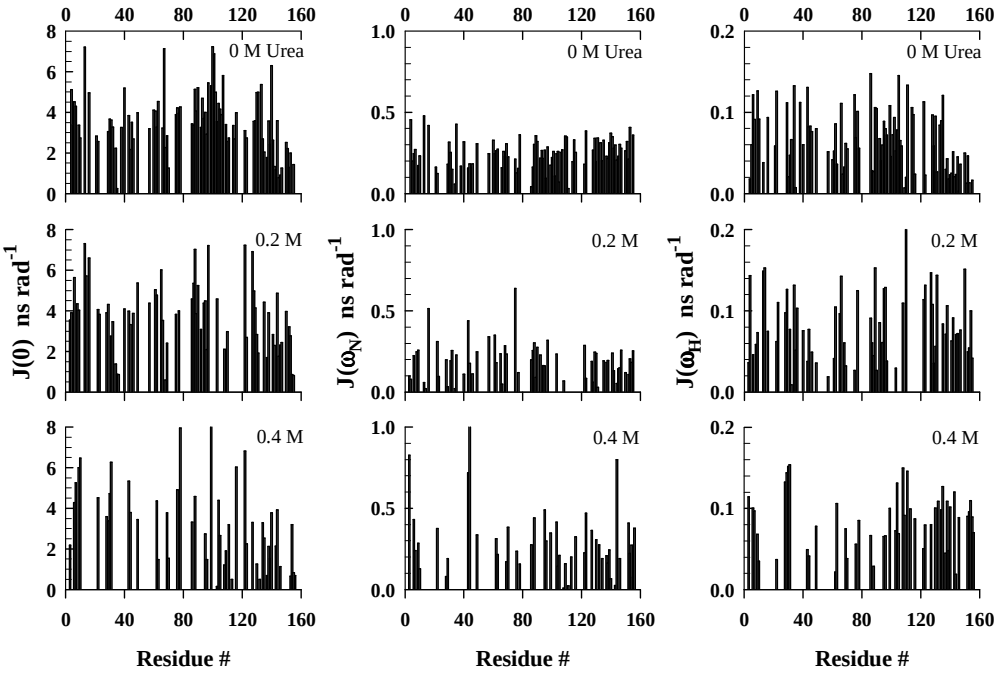


Figure 8. Spectral densities J(0), J(ω_N), and J(ω_H) at indicated concentrations of urea.

$$2\alpha\omega_N^2\tau_i^3+5\beta\omega_N^2\tau_i^2+2(\alpha-1)\tau_i+5\beta=0 \quad (11)$$

where,α and β are slope and intercept, respectively and τ_i can be taken either as τ_c or as τ_e depending on the situation. Excluding the negative roots of τ_i from consideration, the subnanosecond times are assigned to internal correlation times. But, the large values of τ_i do not correspond to tumbling times of molecules the size of *At*PP1; even the value of 20 ns obtained from $J(\omega_0)$ vs $J(\omega_H)$ plot is distantly close to τ_c estimated from T_1:T_2 ratios. One of the reasons for such huge disagreement is corrupt values of J(0) due to pseudo-first order and exchange processes. Hence, the overall tumbling time of *At*PP1 determined from T_1:T_2 ratios (Figure 4) and model-free calculations are more reliable.

3.4.6 Conformational Entropy of *At*PP1. For ps time-scale motion of a NH bond vector such that the vector motion is confined to a cone, the conformational entropy of a bond vector is given by[56]

$$S_{conf} = R \ln\left[\pi\left(3-\sqrt{(1+8S)}\right)\right] \quad (12)$$

The calculated value of S_{conf} for all resolved NH sites of native-state *At*PP1 are positive, averaging to ~2.5 cal mol^{-1} K^{-1} (Figure 10a), indicating intrinsically disordered topology. The values are however sizably low for urea-induced subdenatured states, averaging to -0.31 and 1.11 cal mol^{-1} K^{-1} in the presence of 0.2 and 0.4 M urea, respectively (Figure 10a). Such substantial loss of conformational entropy suggests urea-induced stiffening of the backbone, which is a manifest of the collapse of the intrinsically disordered structure under subdenaturing conditions. The

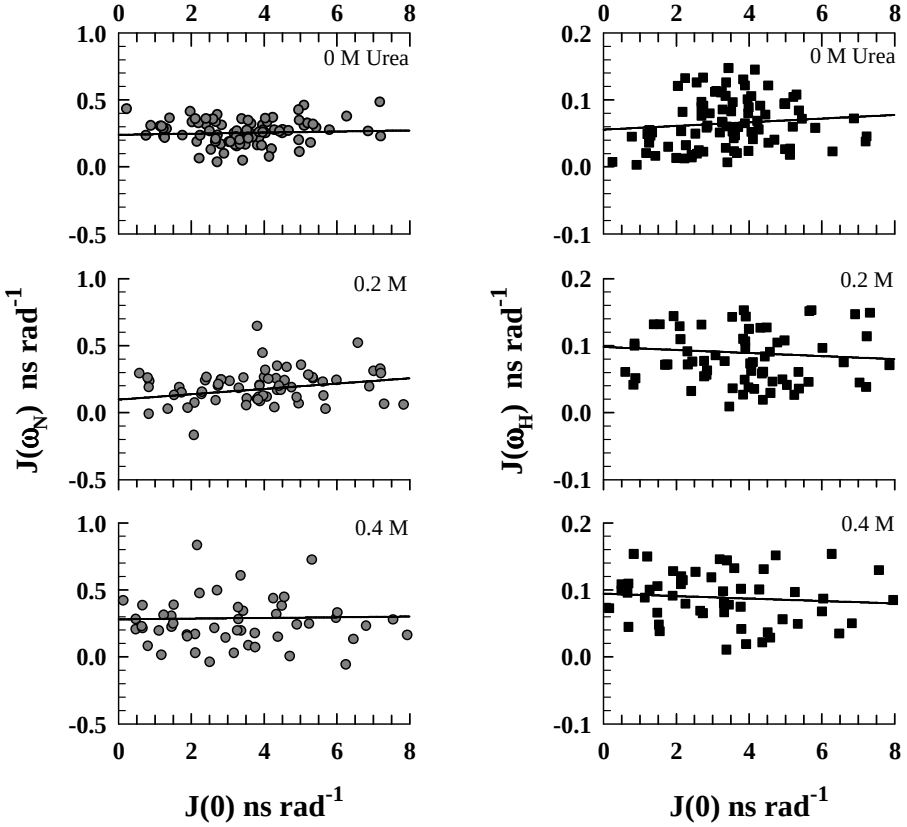


Figure 9. Correlation of $J(\omega_N)$ and J(0), and $J(\omega_H)$ and J(0) at indicated at indicated urea concentrations. Solid lines are linear least square fits.

change in S_{conf} of backbone NHs caused by the addition of urea relative to the S_{conf} in the absence of urea (Figure 10b) ranges from 0.5 to ~6.4 cal mol^{-1} K^{-1} for different residues. These ΔS_{conf} values may be used to extract the backbone conformational free-energy changes at residue level when the native protein is subdenatured in urea

(ΔG=–TΔS_{conf}), and summation of the ΔG values should give the total conformational free energy change ΔG_{conf} associated with the transition of the IDP from the native to urea-induced subdenatured (collapsed) state. However, this provision can be availed only if each NH vector is motionally independent,[57] which is difficult to ascertain. Furthermore, estimation of ΔG_{conf} requires the availability of the ΔS_{conf} value for all residues and under all solvent conditions. Therefore, quantification of ΔG_{conf} at the molecular level needs caution. The results, however, clearly indicate that the subdenatured protein produced by the addition of urea is thermodynamically more stable than that in the absence of urea.

3.5 Discussion

3.5.1 *At*pp1 is an IDP. Whereas the basic experimental results, including the response of the protein fluorescence to denaturant and ANS, provided the initial clue to the content of intrinsically disordered regions (IDRs) in *At*PP1, the compelling evidence for its IDP attribute came from the difficulty of assigning side chain resonances in several stretches of amino acid residues in the course of NMR structure determination. The putative residue segments forming the IDRs in *At*PP1, namely, 17-19, 24-26, 33, 47-57, 70-75, 79-86, 113-131, and 147-149, are in accordance with the idea that the presence of larger number of short disordered regions compared with long ones is a general feature of IDPs.[58] At the experimental pH of 3.6 the protein is highly soluble and stable, and has a net charge of +20 of which 11 are due to basic side chains in the IDRs. The high charge density, which is a characteristic of IDPs,[59] presumably gives rise to random coil-like structures in the IDRs. The finding of disordered regions in the plant protein *At*PP1 is hardly a surprise, because a large

number of cellular proteins - larger in eukaryotes, are more or less intrinsically disordered.[60,61] The important aspects in the IDP context however are their dynamic behavior which is closely related to energetic stability, and the mechanism of disorder-to-collapse transition, both of which have not been studied in detail.

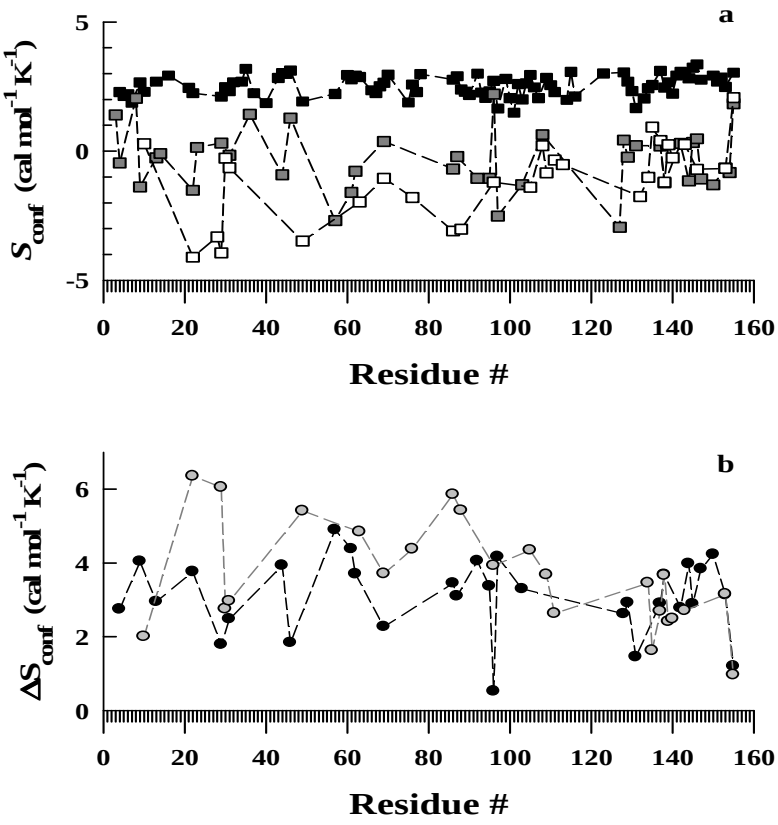


Figure 10. (a) Decrease in conformational entropy, S_{conf}, for various N–H vectors of *At*PP1 in the presence of 0 (black squares), 0.2 (grey squares), and 0.4 M (open squares) urea. (b) Variation of ΔS_{conf}, the difference of S_{conf} between 0 and 0.2 M urea (grey cicrcles), and 0 and 0.4 M urea (open circles).

3.5.2 Backbone Dynamics of *At*PP1. Dynamic disorder and fluctuations in IDRs, the degree of both of which should depend on amino acid types and their relative locations in the disordered segments, have been recognized for a long time.[8,11,62] The conformational fluctuations may often be localized within the disordered regions that serve as target sites for molecular interaction akin to local order-disorder dynamic fluctuations involving active site residues of single-domain proteins.[63] Here, both T_2 and S^2 have shown high degree of widespread dynamic flexibility across the *At*PP1 backbone, not restricted to the putative disordered regions alone. The length of 25 residues at the C-terminus that exhibit larger T_2 and lower J(0) values, both relative to the respective values for the rest of the sequence, is the only region that appears consistently inflexible in both the native and subdenatured states. The segmental inflexibility does not imply rigidity of individual NH vectors, because the S^2 values for these residues is uniformly small which provide no indication for rigidity of the NH sites. This C-terminal segment presumably does not access as many conformational states as other segments do. Since structurally disordered fluctuating regions of IDP are believed to scaffold and interact with other molecules to execute functions,[14,64-69] the C-terminal region of *At*PP1 is probably not involved in the supposed binding with mRNA.

3.5.3 Motional and Structural Collapse of*At*PP1 in Subdenatured States. The dynamic disorder-to-order transition upon placing the protein in increasingly subdenaturing conditions starting from 0 M urea is the central result of this study. The transition entails progressive diminution in polypeptide fluctuations leading to an increasingly compact protein state. The reduction in collective motions of atoms and

fluctuations of chain segments stiffens the structure, which is reminiscent of the collapse of an intrinsically disordered protein to a relatively ordered compact state brought about by restrictions on conformational sampling of the former. This is a chemical denaturant-induced collapse of IDP which may be distinguished from folding coupled to biological ligand binding. In the latter, structurally disordered regions of IDPs mediate interactions with ligand proteins or effector molecules whose binding brings about disorder-to-order transition within them.[64,69,70]

3.5.4 The Mechanism of Urea-Induced Collapse. The basis for urea-induced collapse of IDP, indeed proteins in general, is that the denaturant serves as an honorary ligand. The mechanism of action of chemical denaturants on proteins might appear controversial, because evidences exist for both direct effect exerted by binding of denaturants to the protein,[71-78] and indirect effect caused by shifting the $(H_2O)_{bulky} \leftrightarrow (H_2O)_{dense}$ equilibrium toward the right.[79,80,81] It might as well as be that both direct and indirect effects operate in the denaturant-containing protein solution.[82-87] Whether or not chemical denaturants alter water structure and surface hydration, experimental evidences based on isothermal calorimetric study[88] and x-ray crystallography[89-91] are compelling to believe that they directly bind to proteins. The nature of binding with native-like or subdenatured proteins often involves polyfunctional interactions of denaturants with protein side-chain and backbone atoms.[72,88,91] Van der Waals and dispersive interactions,[91,92] and multiple hydrogen bonding interactions between denaturant and protein[90,91] have been described. Chemical denaturants can thus be regarded an honorary ligands which do not require a particular pocket or target site on a protein to bind to. When used in low

concentrations, their widespread networking interactions effectively cross-link different parts of the molecule leading to a subdenatured state which is dynamically and conformationally collapsed. The denaturant-induced collapsed state is stiff and hydrodynamically shrunk,[33,93,94] and is smothered in terms of entropy, but may gain thermodynamic stability due to the interactions established with the denaturant.[74,75] It is important to note that such a collapse is defined only when the denaturant level used is low and within the subdenaturing limit. At higher concentrations, destabilizing and eventual unfolding effect of the denaturant will overrun its own initial deed of collapse.

The protein stiffening and collapse discussed here is a completely general effect. It becomes distinct and is amplified in the case of IDPs because of the occurrence of a high degree of dynamic restriction. Using some of the typical thermodynamic and binding parameters for single-domain protein-denaturant interactions published by Makhatadze and Privalov,[88] it is estimated that two to three urea molecules would bind to each molecule of *At*PP1 under experimental conditions of 0.2 or 0.4 M urea employed in this study. This estimate may vary because the flexibility of disordered binding regions of IDP may allow for broad specificity, but the binding affinity is reduced by a loss of conformational entropy (Figure 10) due to the ensuing collapse.[9,95] The collapse protein is also expected to have slower internal diffusion arising from entropy loss and increased restraints on non-bonded atom repulsions.[94,96]

3.5.5 The Mechanism of GdnHCl-Induced Collapse. A relatively recent study reported IDP collapse in the presence of low concentrations of GdnHCl with

the correct interpretation that $GdnH^+$ cations shield the negative charges on the IDPs,[33] which is consistent with the idea that expansion of IDP chains and their hydrodynamic dimensions depend on the mean net charge content.[29,97] However, from the preceding discussion on the bindning of chemical denaturants to proteins it is clear that the effect of GdnHCl is produced not only by charge screening action of $GdnH^+$, but also the direct binding interactions that stiffen and shrink the protein globally. Hence, the interpretations of the results of Müller-Späth et al.[33] must also include the direct effect of $GdnH^+$ binding that would contribute to IDP collapse.

3.5.6 Protein Unfolding by Chemical Denaturants. The dynamic and structural collapse of proteins under subdenaturing conditions discussed here is a general phenomenon as far as chemical denaturants are employed to study the unfolding reaction, and bears obvious relevance to structural and thermodynamic properties of end-state ensembles across the unfolding equilibrium (N↔U). To recapitulate the idea, increments of subdenaturing amount of chemical denaturants takes the protein through a large number of native-like or subdenatured states which are increasingly constrained both locally and globally. These constraints cease to occur in the presence of unfolding levels of the denaturant, because binding of a large number of denaturants unfolds the structure globally. The unfolded chain then continues to expand because additional denaturant binding increases the chain volume. These changes often escape detection, even with highly sensitive monitoring probes, and the resulting unfolding transition is modeled as a cooperative two-state transition. The insufficiency of such analysis can be easily checked by multiple perturbation experiments that employ more than one denaturing variable.[98-101] Results

from such studies now provide compelling evidence that native-state free energy surfaces of proteins are vastly rough consisting of numerous subdenatured states[102-104] that are alike in chain topology and secondary structure content, but are distinct in terms of tertiary structure, dynamic flexibility, and average molecular dimension.[93,94,100]

3.6 References

1. Huber, R. Conformational Flexibility and its Functional Significance in Some Protein Molecules. *TrendsBiochem*. *Sci*. **1979**, *4*, 271-276.
2. Riek, R.; Hornemann, S.; Wider, G.; Billeter, M.; Glockshuber, R.; Wuthrich, K. NMR Structure of the Mouse Prion Protein Domain PrP(121-321). *Nature***1996**, *382*, 180-182.
3. Muchmore, S. W.; Sattler, M; Liang, H.; Meadows, R. P.; Harlan, J. E.; Yoon, H. S; Nettesheim, D.; Chang, B.S.; Thompson, C. B.; Wong, S. L.; Ng, S. L.; Fesik, S. W. X-ray and NMR Structure of Human Bcl-x_L, an Inhibitor of Programed Cell Death. *Nature***1996**, *381*, 335-341.
4. Kriwacki, R.W.; Hongst, L.; Tennant, L.; Reed, S. I.; Wright, P. E. Structural Studies of $p21^{Waf1/CiP1/Sdi1}$ in the Free and Cdk2-Bound State: Conformational Disorder Mediates Binding Diversity. *Proc*. *Natl*. *Sci*. *Acad*. *USA***1996**, *93*, 11504-11509.
5. Fletcher, C. M.; Wagner, G. The Interaction of eIF4E with 4E-BP1 is an Induced Fit to a Completely Disordered Protein. *ProteinSci*. **1998**, *7*, 1639-1642.
6. Wright, P. E.; Dyson, H. J. Intrinsically Unstructured Proteins: Re-assessing the Protein Structure-Function Paradigm. *J*. *Mol*. *Biol*. **1999**, *293*, 321-331.
7. Dyson, H. J.; Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem*. *Rev*. **2004**, *104*, 3607-3622.
8. Mittag, T.; Forman-Kay, J. D. Atomic-Level Characterization of Disordered Protein Ensembles. *Curr*. *Opin*. *Struct*. *Biol*. **2007**, *17*, 3-14.
9. Wright, P. E.; Dyson, H. J. Linking Folding and Binding. *Curr*. *Opin*. *Struct*. *Biol*. **2009**, *19*, 31-38.

10. Eliezer, D. Biophysical Characterization of Intrinsically Disordered Proteins. *Curr*. *Opin*. *Struct*. *Biol*. **2009**, *19*, 23-30.
11. Mittag, T.; Marsh, J.; Grishaev, A.; Orlicky, S.; Lin, H.; Sicheri, F.; Tyers, M.; Forman-Kay, J. D. Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with the Cdc4 Subunit of an SCF Ubiquitin Ligase. *Structure***2010**, *18*, 494-506.
12. Kosol, S.; Contreras-Martos, S.; Cedeno, C.; Tompa, P. Structural Characterization of Intrinsically Disordered Proteins by NMR Spectroscopy. *Molecules***2013**, *18*, 10802-10828.
13. Felli, I.; Connelli, L.; Pierattelli, R. In-Cell ^{13}C NMR Spetroscopy for the Study of Intrinsically Disordered Proteins. *Nat*. *Protocols***2014**, *9*, 2005-2016.
14. Van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of Intrinsically Disordered Regions and Proteins. *Chem*. *Rev*. **2014**, *114*, 6589-6631.
15. Daughdrill, G. W.; Chadsey, M. S.; Karlinsey, J. E.; Hughes, K. T.; Dahlquist, F. W. The C-Terminal Half of the Anti-Sigma Factor, FlgM, becomes Structured When Bound to its Target, Sigma 28. *Nat*. *Strut*. *Biol*. **1997**, *4*, 285-291.
16. Breidenbach, M. A.; Brunger, A. T. Substrate Recognition Strategy for Butulinum Neurotoxin Seritype A. *Nature***2004**, *432*, 925-929.
17. Mohan, A.; Oldfield, C. J.; Radivojac, P.; Vacic, V.; Cortese, M. S.; Dunker, A. K.; Uversky, V. N. Analysis of Molecualr Recognition Features (MoRFs). *J*. *Mol*. *Biol*. **2006**, *362*, 1043-1059.

18. De Guzman, R. N.; Goto, N. K.; Dyson, H. J.; Wright, P. E. Structural Basis for Cooperative Transcription Factor Binding to the CBP Coactivator. *J. Mol. Biol.* **2006**, *355*, 1005-1013.
19. Sugase, K.; Dyson, H. J.; wright, P. E. Mechanism of Coupled Folding and Binding of an Intrinsically Disordered Protein. *Nature***2007**, *447*, 1021-1025.
20. Dogan, J.; Gianni, S.; Jemth, P. The Binding Mechanisms of Intrinsically Disordered Proteins. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6323-6331.
21. Sigalov, A. B.; Zhuravleva, A. V.; Orekhov, V. Y. Binding of Intrinsically Disordered Proteins is not Necessarily Accompanied by a Structural Transition to a Folded Form. *Biochimie***2007**, *89*, 419-421.
22. Tompa, P.; Fuxreiter, M. Fuzzy Complexes: Polymorphism and Structural Disorder in Protein-Protein Interactions. *TrendsBiochem. Sci.* **2008**, *33*, 2-8.
23. Hazy, E.; Tompa, P. Limitations of Induced Folding in Molecular Recognition by Intrinsically Disordered Proteins. *Chem. Phys. Chem.* **2009**, *10*, 1415-1419.
24. Foman-Kay, J. D.; Mittag, S. From Sequence and Forces to Structure, Function, and Evolution of Intrinsically Disordered Proteins. *Structure***2013**, *21*, 1492-1499.
25. Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence Correlation Spectroscopy Shows that Monomeric Polyglutamine Molecules form Collapsed Structures in Aqueous Solutions. *Proc. Natl. Sci. Acad. USA***2006**, *103*, 16764-16769.
26. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. Natively Unfolded Yeast Prion Monomer Adopts and Ensemble of Collapsed and Rapidly Flutuating Structures. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 2649-2654.
27. Walters, R.; Murphy, R. Examining polyglutamine peptide length: A Connection

Between Collapsed Conformations and Increased Aggregation. *J. Mol. Biol.* **2009**, *393*, 978-992.

28. Dougan, L.; Li, J.; Badilla, C.; Berne, B. J.; Fernandez, J. Single Homopolypeptide Chains Collapse into Mechanically Rigid Conformations. *Proc. Natl. Acad. Sci. USA***2009**, *106*, 12605-12810.
29. Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net Charge Per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. USA***2010**, *107*, 8183-8188.
30. Gambin, Y.; Vandelinder, V.; Ferreon, A. C. M.; Lemke, E. A.; Groisman, A.; Deniz, A. A. Visualizing a One-Way Protein Encounter Complex by Ultrafast Single-Molecule Mixing. *Nat. Methods***2011**, *8*, 239-241.
31. Ferreon, A. C. M.; Deniz, A. A. Protein folding at Single-Molecule Resolution. *Biochim. Biophys. Acta***2011**, *1814*, 1021-1029.
32. Arya, A.; Mukhopadhyay, S. Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein. *J. Phys. Chem.* B **2014**, *118*, 9191-9198.
33. Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuller, B. Charge Interactions can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. USA***2010**, *107*, 14609-14614.
34. Xoconostle-Cazares, B.; Xiang, Y.; Ruiz-Medrano, R.; Wang, H. L.; Monzer, J.; Yoo, B. C.; McFarland, K. C.; Franceschi, V. R.; Lucas, W. J. Plant Paralog to Viral Movement Protein that Potentiates Transport of mRNA into the Phloem. *Science***1999**, *283*, 94-98.

35. Farrow, N.A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L.E. Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 domain Studied by ^{15}N NMR Relaxation. *Biochemistry***1994**, *33*, 5984-6003.
36. Markley, J. L.; Horsley, W. J.; Klein, M. P. Spin-Lattice Relaxation Measurements in Slowly Relaxing Complex Spectra. *J. Chem. Phys*. **1971**, *55*, 3604-3605.
37. Rawitch, A. B.; Hwan, R-Y. Anilinonaphthalene sulfonate as a Probe for the Native Structure of Bovine Alpha Lactalbumin: Absence of Binding to the Native, Monomeric Protein. *Biochem. Biophys. Res. Commun*. **1979**, *91*, 1383-1389.
38. Semisotnov, G. V.; Rodionova, N. A.; Razgulayev, O. I.,; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. Study of the Molten Globule Intermediate State in Protein Folding by a Hydrophobic Fluorescent Probe. *Biopolymers***1991**, *31*, 119-128.
39. Kay, L. E.; Torchina, D. A.; Bax, A. Backbone Dynamics of Proteins as Studied by ^{15}N Inverse Detected Heteronuclear NMR Spectroscopy: Applications to Staphylococcal Nuclease. *Biochemistry***1989**, *28*, 8972-8979.
40. Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. Backbone Dynamics of Calmodulin Studied by ^{15}N Relaxation Using Inverse Detected Two-Dimensional NMR Spectroscopy: The Central Helix is Flexible. *Biochemistry***1992**, *31*, 5269-5278.
41. Fushman, D.; Weisemann, R.; Thuring, H.; Ruterjans, H. Backbone Dynamics of Ribonuclease T1 and its Complex with 2′GMP Studied by Two-Dimensional Heteronuclear NMR Spectroscopy. *J. Biomol. NMR***1994**, *4*, 61-78.
42. Abragam, A. *PrinciplesofNuclearMagnetism*; Clarendon Press: Oxford, 1961.
43. Lipari, G.; Szabo, A. Model-free Approach to the Interpretation of Nuclear

Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **1982**, *104*, 4546-4559.

44. Lipari, G.; Szabo, A. Model-free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *J. Am. Chem. Soc.* **1982**, *104*, 4559-4570.
45. Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. *J. Am. Chem. Soc.* **1990**, *112*, 4989-4991.
46. Palmer, A. G.; Rance, M.; Wright, P. E. Intramolecular Motions of a Zinc Finger DNA-Binding Domain from Xfin Chracterized by Proton-detected Natural Abundance C-12 Heteronuclear NMR Spectroscopy. *J. Am. Chem. Soc.* **1991**, *113*, 4371-4380.
47. Mandel, A. M.; Akke, M.; Palmer, A. G. Backbone Dynamics of Escherichia Coli Ribonuclease H1 – Correlations with Structure and Function in an Active Enzyme. *J. Mol. Biol.* **1995**, *246*, 144-163.
48. Lefevre, J.-F.; Dayie, K. T.; Peng, J. W.; Wagner, G. Internal Mobility in the Partially Folded DNA Binding and Dimerization Domains of GAL4: NMR Analysis of the N-H Spectral Density Functions. *Biochemistry***1996**, *35*, 2674-2686.
49. Rundqvist, L.; Tengel, T.; Zdunek, J.; Bjorn, E.; Schleucher, J.; Alcocer, M. J.C.; Larsson, G. Solution Structure, Copper Binding and Backbone Dynamics of Recombinant Ber e 1 – The major Allergen from Brazil Nut. *PLoSONE***2012**, *7*, e46435.

50. Hiyama, Y.; Niu, C-H.; Silverton, J. V.; Bavoso, A.; Torchia, D. A. Determination of ^{15}N Chemical Shift Tensor via ^{15}N-^{2}H Dipolar Coupling in Boc-Glycylglycyl[^{15}N]glycine Benzyl Ester. *J. Am. Chem. Soc*. **1988**, *110*, 2378-2383.
51. Peng, J. W.; Wagner, G. Mapping of Spectral Density Functions Using Heteronuclear NMR Relaxation Mesurements. *J. Mag. Reson*. **1992**, *98*, 308-332.
52. Peng, J. W.; Wagner, G. Mapping of the Spectral Densities of N-H Bond Motions in Eglin C Using Heteronuclear Relaxation Measurements. *Biochemistry***1992**, *31*, 8571-8586.
53. Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. Comparison of the Backbone Dynamics of a Folded and an Unfolded SH3 Domain Existing in Equilibrium in Aqueous Bufer. *Biochemistry***1995**, *34*, 868-878.
54. Peng, J. W.; Wagner, G. Investigation of Protein Motions via Relaxation Measurements. *MethodsEnzymol*. **1994**, *239*, 563-596.
55. Zhang, P.; Dayie, K. T.; Wagner, G. Unusual Lack of Internal Mobility and Fast Overall Tumbling in Oxidized Flavodoxin from *Anacystisnidulans*. *J. Mol. Biol*. **1997**, *272*, 443-455.
56. Yang, D.; Kay, L. E. Contributions to Conformational Entropy Arising from Bond Vector Fluctuations Measured from NMR-Derived Order Parameters: Application to Protein Folding. *J. Mol. Biol*. **1996**, *263*, 369-382.
57. Sarata C. Sahu;Abani K. Bhuyan;Ananya Majumdar and Jayant B. Udgaonkar. Backbone Dynamics of Barstar: A ^{15}NNMRRelaxation Study. *Proteins* 2000, 41, 460–474
58. Xue, B.; Williams, R. W.; Oldfield, C. J.; Goh, G. K.; Dunker, A. K.; Uversky, V. N. Viral Disorder or Disordered Viruses: Do Viral Proteins Possess Unique

Features? *ProteinPept. Lett.* **2010**, *17*, 932.

59. Uversky, V. N. Natively Unfolded Proteins: A Point where Biology Waits for Physics. *ProteinSci.* **2002**, *11*, 739-756.
60. Dunker, A. K.; Brown, C. J.; Lawson, D.; Iakoucheva, L. M.; Obradovic, Z. Intrinsic Disorder and Protein Function. *Biochemistry***2002**, *41*, 6573-6582.
61. Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **2004**, *337*, 635-645.
62. Dyson, H. J.; Wright, P. E. Nuclear Magnetic Resonance Methods for Elucidation of Structure and Dynamics in Disordered States. *MethodsEnzymol.* **2001**, *339*, 258-270.
63. Whitten, S. T.; Garcia-Moreno, E. B.; Hilser, V. J. Local Conformational Fluctuations can Modulate the Coupling Between Proton Binding and Global Structural Transitions in Proteins. *Proc. Natl. Acad. Sci. USA***2005**, *102*, 4282-4287.
64. Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and their Functions. *Nat. Rev. Mol. CellBiol.* **2005**, *6*, 197-208.
65. Meszaros, B.; Tompa, P.; Simon, I.; Dosztanyi, Z. Molecular Principles of the Interactions of Disordered Proteins. *J. Mol. Biol.* **2007**, *372*, 549-561.
66. Bertagna, A.; Toptygin, D; Brand, L.; Barrick, D. The Effects of Conformational Heterogeneity on the Binding of the Notch Intracellular Domain to Effector Proteins: A Case of Biologically Tuned Disorder. *Biochem. Soc. Trans.* **2008**, *36*, 157-166.
67. Mittag, T.; Orlicky, S.; Choy, W.; Tang, X.; Lin, H.; Sicheri, F.; Kay, L. E.; Tyers,

M.; Forman-Kay, J. D. Dynamic Equilibrium Engagement of a Polyvalent Ligand with a Single Site Receptor. *Proc. Natl. Acad. Sci. USA***2008**, *105*, 17772-17777.

68. Ganguly, D.; Chen, J. Structural Interpretation of Paramagnetic Relaxation Enhancement-Derived Distances for Disordered Protein States. *J. Mol. Biol.* **2009**, *390*, 467-477.
69. Babu, M. M.; Van der Lee, R.; Sanchez de Groot, N.; Gsponer, J. Intrinsically Disordered Proteins: Regulation and Disease. *Curr. Opin. Struct. Biol.* **2011**, *21*, 1-9.
70. Borg, M.; Mittag, T.; Pawson, T.; Tyers, M.; Forman-Kay, J. D.; Chan, H. S. Polyelectrostatic Interactions of Disordered Ligands Suggest a Physical Basis for Ultrasensitivity. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 9650-9655.
71. Simpson, R. B.; Kauzmann, W. The Kinetics of Protein Denaturation. I. The Behavior of the Optical Rotation of Ovalbumin in Urea Solutions. *J. Am. Chem. Soc.* **1953**, *75*, 5139-5152.
72. Robinson, D. R.; Jencks, W. P. The Effect of Compounds of the Urea-Guanidinium Class on the Activity coefficient of Acetyltetraglycine Ethyl Ester and Related Compounds. *J. Am. Chem. Soc.* **1965**, *87*, 2462-2470.
73. Breslow, R.; Guo, T. Surface Tension Measurements Show that Chaotropic Salting-in Denaturants are Not Just Water-Structure Breakers. *Proc. Natl. Acad. Sci. USA***1990**, *87*, 167-169.
74. Bhuyan, A. K. Protein Stabilization by Urea and Guanidine Hydrochloride. *Biochemistry***2002**, *41*, 13386-13394.
75. Tobi, D.; Iber, R.; Thirumalai, D. The Dominant Interaction Between Peptide and Urea is Electrostatci in Nature: A Molecualr Dynamics Simulation Study.

*Biopolymers***2003**, *68*, 359-369.

76. Kumar, R.; Prabhu, N. P.; Yadaiah, M.; Bhuyan, A. K. Protein Stiffening and Entropic Stabilization in the Subdenaturing Limit of Guandidine Hydrochloride. *Biophys*. *J*. **2004**, *87*, 2656-2662.
77. Street, T. O.; Bolen, D. W.; Rose, G. D. A Molecular Mechanism for Osmolyte-Induced Protein Stability. *Proc*. *Natl*. *Acad*. *Sci*. *USA***2006**, *103*, 13997-14002.
78. O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. Interactions Between Hydrophobic and Ionic Solutes in Aqueous Guanidinium Chloride and Urea Solutions: Lessons from Protein Denaturation Mechanisms. *J*. *Am*. *Chem*. *Soc*. **2007**, *129*, 7346-7353.
79. Frank, H. S.; Evans, M. W. Free Volume and Entropy in Condensed Systems. III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J*. *Chem*. *Phys*. **1945**, *13*, 507-532.
80. Finer, E. G.; Franks, F.; Tait, M. J. Nuclear Magnetic Resonance Studies of Aqueous Urea Solutions. *J*. *Am*. *Chem*. *Soc*. **1972**, *94*, 4424-4429.
81. Muller, N. Search for a Realistic View of Hydrophobic Effects. *Acc*. *Chem*. *Res*. **1990**, *23*, 23-28.
82. TiradoRives, J.; Orozco, M.; Jorgensen, W. L. Molecular Dynamics Simulations of the Unfolding of Barnase in Water and 8 M Aqueous Urea. *Biochemistry***1997**, *36*, 7313-7329.
83. Caflisch, A.; Karplus, M. Structural Details of Urea Binding to Barnase: A Molecular Dynamics Analysis. *Structure***1999**, *7*, 477-488.
84. Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Barnes, A. C.; Cruickshank, J. M.

The Hydration Structure of Guanidinium and Thiocyanate Ions: Implications for Protein Stability in Aqueous Solutions. *Proc. Natl. Acad. Sci. USA***2003**, *100*, 4557-4561.

85. Bennion, B. J.; Daggett, V. The Molecular Basis for the Chemical Denaturation of Proteins by Urea. *Proc. Natl. Acad. Sci. USA***2003**, *100*, 5142-5147.
86. Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L. Effect of Urea on Peptide Conformation in Water: Molecular Dynamics and Experimental Characterization. *Biophys. J.* **2005**, *89*, 842-857.
87. Stumpe, M. C.; Grubmuller, H. Interaction of Urea with Amino Acids: Implications for Urea-Induced Protein Denaturation. *J. Am. Chem. Soc.* **2007**, *129*, 16126-16131.
88. Makhatadze, G. I.; Privalov, P. L. Protein Interactions with Urea and Guanidinium Chloride. A Calorimetric Study. *J. Mol. Biol.* **1992**, *226*, 491-505.
89. Hibbard, L. S.; Tulinsky, A. Expression of Functionality of α-Chymotrypsin. Effects of Guanidine Hydrochloride and Urea in the Onset of Denaturation. *Biochemistry***1978**, *17*, 5460-5468.
90. Pike, A. C. W.; Acharya, R. A structural Basis for the Interaction of Urea with Lysozyme. *ProteinSci.* **1994**, *3*, 706-710.
91. Dunbar, J.; Yennawar, H. P.; Banerjee, S.; Luo, J.; and Farber, G. The Effect of Denaturants on Protein Structure. *ProteinSci.* **1997**, *6*, 1727-1733.
92. Hua, L.; Zhou, R.; Thirumalai, D.; Berne, B. J. Urea Denaturation by Stronger Dispersion Interactions with Proteins than Water Implies a 2-Stage Unfolding. *Proc. Natl. Acad. Sci. USA***2008**, *105*, 16928-16933.
93. Rao, M. T.; Bhuyan, A. K.; Venu, K.; Sastry, V. S. S. Nonlinear Effect of GdnHCl

on Hydration Dynamics of Proteins: a 1H Magnetic Relaxation Dispersion Study. *J. Phys. Chem. B***2009**, *113*, 6994-7002.

94. Yasin, U. M.; Sashi, P.; Bhuyan, A. K. Expansion and Internal Friction in Unfolded Protein Chain. *J. Phys. Chem. B***2013**, *117*, 12059-12064.
95. Rader, S. D.; Hellinga, H. W.; Spicer, L. D. Conformatinal Substates in Enzyme Mechansim: The 120K Structure of α-Lytic Protease at 1.5Å Resolution. *ProteinSci*. **1997**, *6*, 1375-1386.
96. Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely Slow Intramolecular Diffusion in Unfolded Protein L. *Proc. Natl. Acad. Sci. USA***2010**, *107*, 13713-13717.
97. Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "Natively Unfolded" Proteins Unstructured Under Physiologic Conditions? *Proteins***2000**, *41*, 415-427.
98. Gupta, R.; Yadav. S.; Ahmad, F. Protein Stability: Urea-induced versus Guanidine-induced Unfolding of Metmyoglobin. *Biochemistry***1996**, *35*, 11925-11930.
99. Gupta, R.; Ahmad, F. Protein Stability: Functional Dependence of Denaturational Gibbs Energy on Urea Concentration. *Biochemistry***1999**, *38*, 2471-2479.
100. Ferreon, A. C. M.; Bolen, D. W. Thermodynamics of Denaturant-induced Unfolding of a Protein that Exhibits Variable Two-state Denaturation. *Biochemistry***2004**, *43*, 13357-13369.
101. Oliva, F. Y.; Munõz, V. A Simple Thermodynamic Test to Discriminate Between Two-state and Downhill Folding. *J. Am. Chem. Soc*. **2004**, *126*, 8596-8597.
102. Campos, L. A.; Sadqi, M.; Liu, J.; Wang, X.; English, D. S.; Munõz, V. Gradual Disordering of the Native State on a Slow Two-state Folding Protein Monitored by

Single-molecule Fluorescence Spectroscopy and NMR. *J. Phys. Chem. B117*, 13120-13131.

103. Naganathan, A. N.; Perez-Jimenes, R.; Sanchez-Ruiz, J. M.; Munõz, V. Robustness of Downhill Folding: Guidelines for the Analysis of Equilibrium Folding Experiments on Small Proteins. *Biochemistry***2005**, *44*, 7435-7449.

104. Naganathan, A. N.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Munõz, V. Estimation of Protein Folding Free Energy Barriers from Calorimetric Data by Multi-model Bayesian Analysis. *Phys. Chem. Chem. Phys*.**2011**, *13*, 17064-17076.

Chapter-4

Thermal Effect on Amide Proton Chemical Shifts and Linear Expansion Coefficient of Hydrogen-Bonds in an Intrinsically Disordered Plant Protein, *At*PP1

4.1 Abstract

Although temperature effect on chemical shifts of hydrogen-bonded amide hydrogens in peptides and proteins is widely used to study hydrogen bond geometry, type and stability of secondary structures, and local unfolding processes in proteins, a clear understanding of the magnitude and direction of temperature-influenced shift is lacking. For the *Arabidopsis thaliana* phloem protein 1 (*At*PP1) studied here, chemical shifts of hydrogen-bonded amide proton resonances move linearly downfield with temperature increments in the 293-313 K range. The downfield linear shift is invariable with respect to α-helix and β-sheet structure types, and remains the same even when the protein is placed in subdenaturing conditions by using low levels of urea. Such changes may arise from temperature dependent changes in hydrogen bond geometry. Linear thermal expansion coefficients (LTEC) for the hydrogen bonds are larger by an order of magnitude when compared to those for other proteins, likely due to intrinsic structural disorder in *At*PP1. The mean LTEC slightly decreases as the native protein is taken to subdenaturing conditions, suggesting some stiffening of structural regions containing urea binding sites. Volume shrinkage and segmental stiffening might be general phenomena for subdenatured proteins.

4.2 Introduction

Hydrogen bonding has been studied with a constant zeal for a little less than hundred years now because of 'something unknown' about the bond and the key role it plays in structure stabilization in chemistry and biology. Emerging ideas, types of hydrogen bonds, and their occurrence in different structures have been discussed from time to time.[1-3] To focus on NMR studies, temperature dependence of chemical shifts (δ_H) of amide proton ($^1H^N$) resonances in peptides and proteins have often served in the past as a marker of hydrogen bonding status of residues,[4-6] although the generality of this marker has been questioned.[7] In fact, the response of δ_H of hydrogen-bonded amide protons to temperature and solvent conditions has been difficult to understand, even though the general idea for shift interpretation is based on thermally driven distortion of the hydrogen bond geometry, direction of the lone pair orbitals on the acceptor oxygen, and waning of the hydrogen bond length. A new idea came to the fore with the description of $^{3h}J_{NC'}$ scalar interaction in N−H···O=C of proteins.[8-10] The observation of hydrogen bond-mediated scalar interactions not only associates covalency with the bond, but also provides a way for assignment of acceptor carbon resonances. The temperature dependence of the size of $^{3h}J_{NC'}$ has been shown to be extremely useful to derive geometry, thermal expansion coefficient, and chemical shift.[3,8,11,12]

In spite of these advances the reasons for a sizable difference in δ_H of a free amide hydrogen in the random coil state vis-â-vis hydrogen bonded in a secondary structural segment remains poorly understood. A number of factors that determine the electron density around the hydrogen atom and the bond geometry itself appear to work in a complicated manner to produce the observed δ_H. The random-coil δ_H

generally shifts upfield when the $^{1}H^{N}$ is hydrogen bonded in an α-helical segment, and shifts downfield when it becomes a part of a β-sheet structure.[13] Weakening of the hydrogen bond at higher temperature is expected to reverse the shifts, but temperature coefficient of amide hydrogens may actually be a poor indicator of intramolecular hydrogen bonding.[7]

This chapter investigates the thermal effect on chemical shifts of hydrogen bonded amide hydrogens of the phloem protein 1 from *Arabidopsis thaliana*, called *At*PP1, involved in a variety of functions, including calcium binding and nutrient transport. It is also likely to play a role in mRNA delivery, as suggested by studies with phloem protein homologs.[14] Reasons for choosing this protein are the ongoing solution structure studies in this laboratory and the intrinsic structural disorder in the protein. The latter provides an opportunity to examine if disordered tertiary structure is reflected in temperature coefficients and linear thermal expansion coefficient (LTEC) of hydrogen bonded $^{1}H^{N}$.

We also determine the temperature coefficients and LTEC of resolved resonances under subdenaturing conditions by including very low concentrations of urea. The idea here is based on reports that temperature dependent shifts are sensitive to local fluctuations and can often detect accessibility of locally destabilized structural regions to high-energy intermediate-like states.[15,16] The chemical shift for all hydrogen bonded $^{1}H^{N}$ under all conditions are observed to shift downfield exhibiting linear temperature dependence. Values of LTEC for all resonances determined over 283 to 313 K are larger than reported for other proteins. Comparison of LTEC values under different urea concentrations indicate mild stiffening of protein segments as the solvent conditions migrate from native to subdenatuirng.

4.3 Experimental Section

4.3.1 Cloning, Expression and Purification. Recombinant *At*PP1 was overexpressed in BL21(DE3)RIL strain of *E. coli* cells harboring the protein expression vector pET28a(+). For ^{15}N-labeling of amide hydrogens, cells were grown in the standard M9 medium where $^{15}NH_4Cl$ was the sole nitrogen source. Harvested cells were pelleted, suspended in a pH 8 buffer consisting of 20 mM Tris, 50 mM NaCl, and 4 mM imidazole, sonicated, and centrifuged at 15,000 rpm to remove the cell debris. The supernatant was loaded in a nickel affinity column equilibrated in the same buffer, and the protein was eluted by passing 20 mM Tris containing 150 mM imidazole at pH 8. The protein was then dialyzed extensively against 100 mM acetic acid (pH~ 3.5). All steps from cell harvesting to dialysis are carried out at 4°C, and protein purity is ascertained by SDS polyacrylamide gel electrophoresis.

4.3.2 Equilibrium Urea Denaturation Measurements. Variable urea samples were prepared by mixing appropriate volumes of two uniform protein solutions (2 μM), one containing 3 M urea and the other without, both prepared in 50 mM sodium acetate buffer, pH 3.5. Samples were incubated at 25°C for 3-6 hours, and fluorescence emission spectra (excitation at 280 nm) were recorded using a Horiba FluoroMax 4P fluorometer peltier-thermostated at 25°C. The same samples were used to record CD spectra in a AVIV SF420 instrument thermostated at 25°C. Urea dependence of fluorescence intensities at 335 nm and CD ellipticity values at 228 nm, both normalized with respect to the corresponding signals of the native state sample, were used for analyses of the transitions.

4.3.3 Thermal Denaturation Measurements. Temperature denatuartion of *At*PP1 was monitored by both fluorescence and far-UV CD using thermoelectric peltiers. Equilibration time at each temperature was 8 minutes. Reversibility of thermal denaturation was checked by a repeat in which the sample initially heated from 283 to 343 K was cooled to 283 K.

4.3.4 NMR Spectroscopy. ^{1}H-^{15}N HSQC spectra were recorded with ~150 μM protein in 7 mM sodium acetate buffer, pH 3.5, containing variable urea in the 0-0.4 M range. For each concentration of urea, spectra were taken at temperatures in the 283-303 K range with 5 K increment. Reversibility of the protein conformational transition in this range of temperature was confirmed by comparing the 283-K spectra before raising the temperature and after cooling the sample. Typically, 128 experiments each of 32 scans in the ^{15}N dimension were recorded in a 500 MHz Bruker spectrometer (AVIII).

4.4 Results and Discussion

4.4.1 Intrinsically Disordered Structure and Conformational Stability of *At*PP1. Disorder in the structure of *At*PP1 has been evident because of the difficulty of resolving a large number of side-chain and some main-chain resonances in multidimensional NMR experiments that have been performed to determine the solution structure of the protein. This has already been discussed in Chapter 3. For the sake of completeness here, a few relevant results of ANS binding and equilibrium unfolding are presented here as well. The result for the basic experiment of binding of

the dye ANS (8-anilino naphthalene sulfonate) to the native protein is shown in Figure 1a. The idea is the same as mentioned in Chapter 3 that ANS binding would be substantial if the protein is structurally disordered allowing larger accessibility of the dye toward hydrophobic surfaces in the protein interior.[17] Indeed, ANS fluoresces even when no structure-destabilizing agents are present in the solution, suggesting structural disorder in *At*PP1.

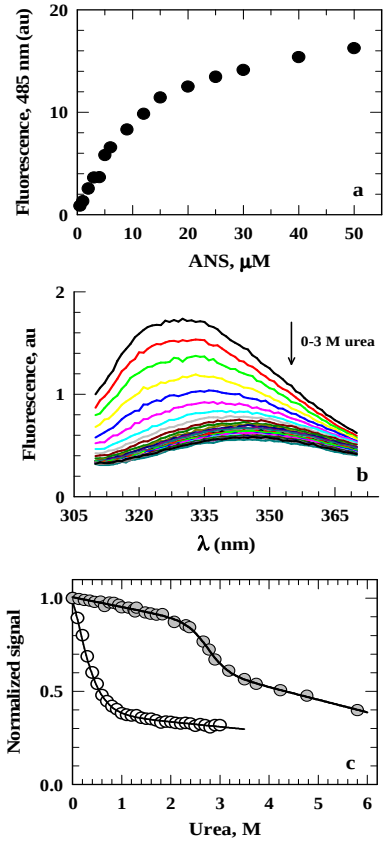


Figure 1. Intrinsic disorder and folding of *At*PP1 at pH 3.5, 25°C. (a) Binding of ANS to the native state suggests the presence of intrinsically disordered structure. Fluorescence was monitored at the emission maximum of ANS. (b) Tryptophan fluorescence emission spectra in 0-3 M range of urea to show quenching of emission, and hence protein unfolding. (c) Urea-induced transition of *At*PP1 monitored by tryptophan fluorescence (open circles) and CD at 228 nm (grey-filled circles). Two-state fit to the CD-monitored data yields $\Delta G°$ values of ~1.3 and 7.5(±0.1) kcal mol^{-1} for unfolding. The solid line through the fluorescence data has been drawn by inspection only, and has no physical meaning.

To examine protein stability, urea denaturation was monitored by intrinsic tryptophan fluorescence and 228-nm CD absorption in 50 mM acetate buffer, pH 3.5. The tertiary structure appears highly unstable and is almost completely lost in ~1 M urea, but the transition for secondary structure occurs only above 2 M urea (Figure 1b,c). As already discussed in Chapter 3, the initial tertiary structure is almost completely removed when urea is added up to ~1 M, which has been interpreted to be an indication of IDP collapse. The transitions are reversibly cooperative, and the CD-monitored global unfolding transition yields a conformational stability of 7.5(±0.1) kcal mol^{-1}. To ensure that the protein does not aggregate and any thermally induced conformational change is reversible, tertiary and secondary structure contents were monitored by successive heating and cooling of the protein in the 283-343 K range of temperature. While the tertiary structure changes reversibly and entirely in this range of temperature (Figure 2a), the secondary structure changes by ~20% (Figure 2b).

4.4.2 Temperature-Dependent Downfield Chemical Shift of Hydrogen Bonded Amide Protons. $^{1}H^{N}$ chemical shifts (δ_H) were measured by ^{1}H-^{15}N HSQC spectra at variable temperature and urea (Figure 3). Thirty three residues whose NH resonances are away from the crowded regions and are identified unambiguously (Figure 3a) were selected for further analysis. Although δ_H values for resonances shift little within the 0-0.4 M range of urea, increasing the temperature from 283 up to 313 K at any urea level invariably produces a linear downfield trend, indicating temperature-induced high-frequency shift or deshielding of hydrogen-bonded amide

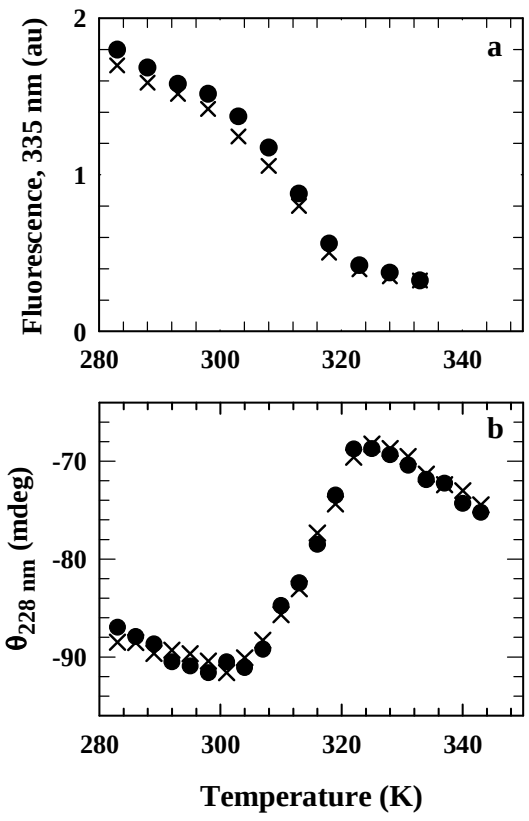
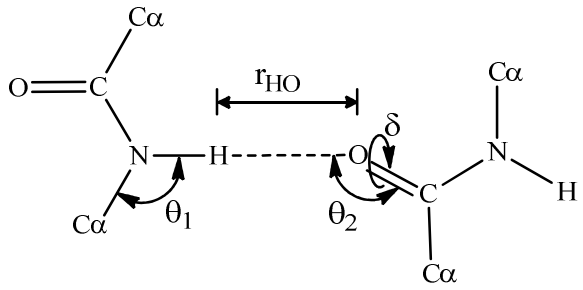


Figure 2. Thermal stability of *At*PP1 at pH 3.5. (a) Fluorescence-monitored reversible thermal denaturation carried out by heating (●) and cooling (×) the sample. (b) Temperature dependence of secondary structure content measured by heating (●) and cooling (×). About 8 minutes were allowed for the sample to equilibrate at each temperature.

protons (Figure 4).The gradient for temperature-dependent chemical shift observed here is positive, which is atypical of the commonly observed negative gradient produced by upfiled shift of δ_H. The shift pattern cannot be related directly to specific structural and conformational transitions, because downfield shifts may arise from a number of factors, including transitions of relatively less structured region to a β-stranded configuration[13] and temperature-dependent excursions of local sites to alternative conformational states.[15] The general understanding is that increasing thermal motions weaken the hydrogen bond by expanding the average bond length, thus leading to greater electron density at the hydrogen and larger shielding generated by the electronic currents of the acceptor carbonyl group. Since thermal motions are not expected to reduce the hydrogen bond

length so as to produce a downfield shift, one may consider temperature effects on the geometry of the bond donor and acceptor shown below.



Earlier work has significantly advanced the understanding of the relation between amide proton shifts and length and geometry of hydrogen bonds.[7,12,18-21] A general dependence of δ_H on the structural parameters r_{OH}, ρ, and θ_2 due to Barfield[12] is given by

$$\delta_H = [4.81\cos^2\theta_2 + \{3.01\cos^2\rho - 0.84\cos\rho + 1.75\}\sin^2\theta_2]\, e^{[-2(r_{OH}-1.76)]} + 4.06 \text{ ppm} \quad (1)$$

according to which the balance of the effects of the three parameters on δ_H determines the effective chemical shift. The positive temperature gradient observed in the present study (Figure 4) would be allowed if the expected upfield shift of δ_H due to larger r_{OH} at higher temperature is overcompensated by departure of either or both of ρ and θ_2.Larger hydrogen bond length at higher temperature presumably increases the transverse vibrations of H···O=C so as to alter the equilibrium value of θ_2.Temperature-induced changes in the dihedral angle, ρ, which has its origin in the protein conformation, can affect the δ_H values by polarizing the electron density and

changing the N−H distance.[21]

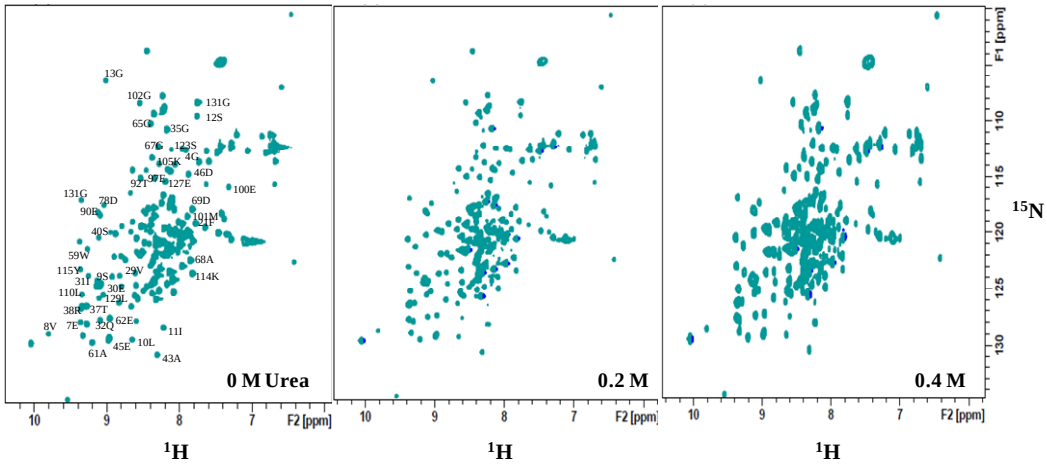


Figure 3. ^{15}N-1H HSQC spectrum of *At*pp1 at pH 3.5 and 298 K, at the urea concentrations indicated. Only labeled resonances were used for chemical shift measurements.

To check if temperature affects the secondary structure of *At*PP1, CD measurements were carried out within the range of temperature and urea that were employed to measure δ_H. For all concentrations of urea, the ellipticity at 228 nm becomes more negative as the temperature is raised from 283 to 298 K (Figure 5), suggesting a gain in the β-sheet character, but gradual loss of ellipticity at higher temperatures indicates β-sheet weakening. Temperature is known to cause such changes and redistribution of secondary structural element in intrinsically disordered proteins.[22] Since *At*PP1 belongs to structurally disordered class of proteins, the observed thermal effects may be related to changes in secondary structure and conformation, although data are insufficient to tell if such changes are due to continuous variation of θ_2 and ρ. The resident sites of these hydrogen bonds are

apparently highly stable both structurally and energetically so that the secondary structural parameters are at least preserved in the entire temperature range. Since chemical shifts and temperature-dependent trends vary minimally at low concentrations of urea (Figure 3), any thermally driven structural alteration which may produce downfield shifts of $^{1}H^{N}$ resonances is not affected under subdentauring conditions.

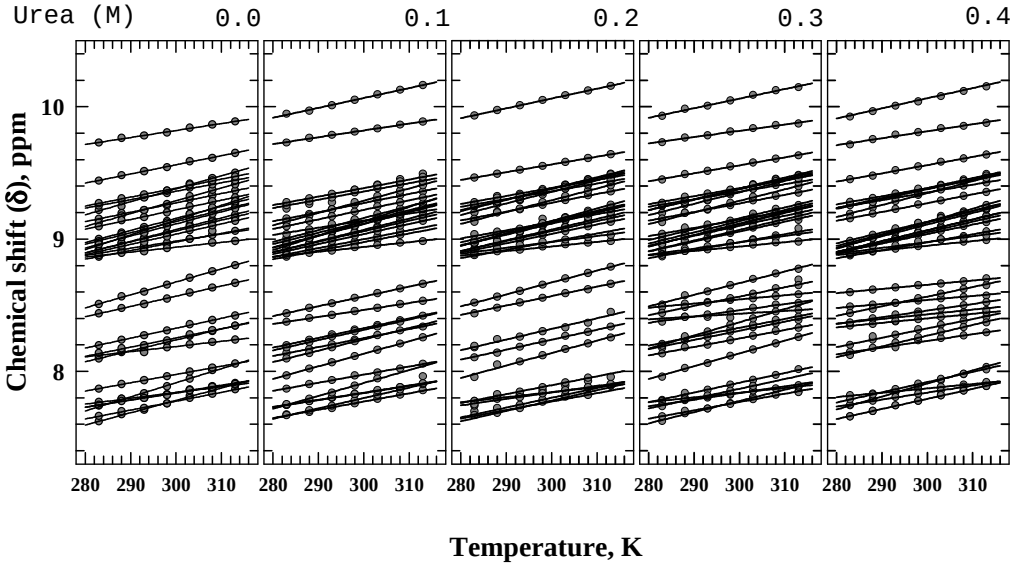


Figure 4. Temperature dependence of chemical shifts of some of the resolved hydrogen-bonded amide protons at the indicated concentrations of urea, pH 3.5. Thermal reversibility of peak shifts was checked for each sample. The resonances correspond to those labeled in Figure 3.

The above discussion is defensive of hydrogen bond weakening due to thermal motions in spite of downfield shifts of δ_H for all hydrogen bonds. In the case

of ubiquitin however, where a β-sheet $^{1}H^{N}$ exhibits such a downfield shift, the authors deliberate that higher temperature truly reduces the length of this hydrogen bond, and hence strengthens it.[11] In the present study, all $^{1}H^{N}$ irrespective of their location in secondary structure type show the downfield shift. Should this mean a negative thermal expansion for all β-sheet hydrogen bonds rather than the speculations of temperature dependent changes in bond geometry and conformation remains open to investigation.

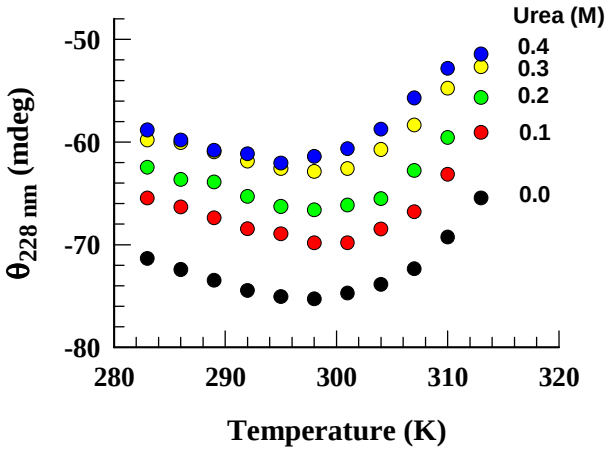


Figure 5. Effect of 283-313 K range of temperature on changes in secondary structure distribution / adjustments at indicated concentrate-ons of urea, pH 3.5. The changes are reversible.

4.4.3 Hydrogen Bond Length and Thermal Expansion Coefficient. Even for downfield shift of δ_H, the arguments above maintain the basic tenet that thermal motions increase the length r_{OH}. Simple approaches have been provided to calculate the average value of r_{OH} when the geometric parameters ρ and θ_2, and/or the trans H-

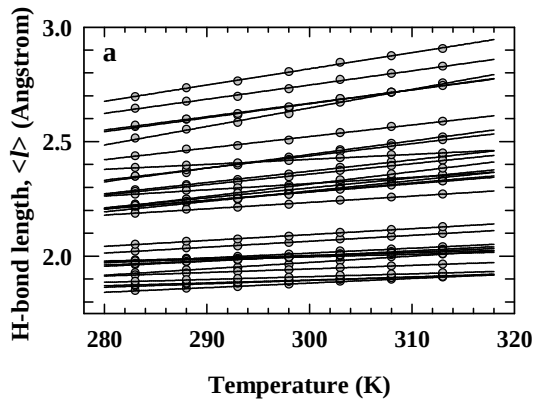
bond coupling ($^{3h}J_{NC'}$) are known.[10,12,20,23] We did not determine the $^{3h}J_{NC'}$ values, because highly accurate measurements are warranted to derive meaningful correlations. Hence, the following linear relation between δ_H and r_{OH} due to Wagner et al[18]

$$\Delta\delta_H = \frac{19.2}{r_{OH}^3} - 2.3 \quad (2)$$

was used to determine the hydrogen bond distances. The $\Delta\delta_H$ value is obtained by subtracting the corresponding random coil shift from the observed chemical shift. Figure 6a shows the increase in average r_{OH} as the protein is heated from 283 to 313 K, from which the linear thermal expansion coefficient of the amide hydrogen bond can be determined by the relation

$$\alpha(T_1, T_2) = \frac{1}{\langle r_{OH} \rangle} \frac{\partial r_{OH}}{\partial T} \quad (3)$$

where,$<r_{OH}>\approx$2.1 Å is the average hydrogen bond length.Residue distribution of α over the temperatures from 283 to 313 K for the native state of the protein (Figure 6b) shows a general trend of higher values for hydrogen bonds in β-sheet regions. The mean values, $<\alpha>$, are 2.204×10^{-3} and 2.422×10^{-3} K^{-1} for hydrogen bonds in all structural regions and β-sheets alone, respectively. The $<\alpha>$ value determined for *At*PP1 here is about an order of magnitude larger than those for proteins ubiquitin[11] and GB3,[12] all measured within comparable temperature ranges. This reflects weaker hydrogen bonding in *At*PP1 which could be due to intrinsically disordered structure of the protein.



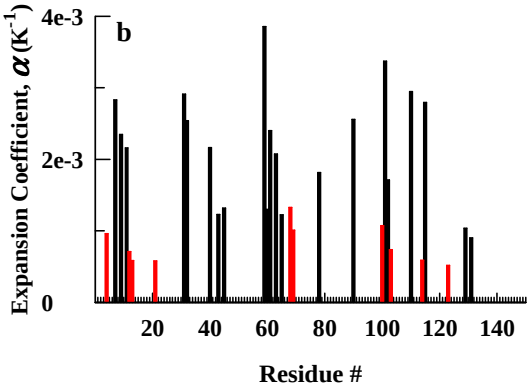


Figure 6. Influence of temperature on amide hydrogen bond expansion in native *At*PP1, pH 3.5. (a) Increase in hydrogen bond length with temperature calculated by applying the relation of Wagner et al (18)(ref) to measured data. (b) Linear thermal expansion coefficient for α-helical and β-sheet residues (red and black bars, respectively).

4.4.4 Variation of α Under Subdenaturing Conditions. To examine thermal expansion of the hydrogen bonds in the pretransition region of the unfolding transition Figure 7a shows the range of α values for all hydrogen bonds at different levels of urea in the 0-0.4 M range, where the spread in α (~3.5×10^{-4} to ~3×10^{-3} K^{-1}) seems comparable from one urea concentration to another. But the mean of the spread, shown for hydrogen bonds in all regions and those in β-sheet segments (Figure 7b), indicates a slight decrease in the value of $<\alpha>$ as the urea level is increased from 0 to 0.2 M. The decrease from 2.422×10^{-3} to 1.713×10^{-3} K^{-1} for β-sheet hydrogen bonds, for example, appears marginal, but circumstantial evidences

exist to suggest that proteins might actually shrink in the presence of subdenaturing levels of urea and guanidinium hydrochloride (GdnHCl). Polyfunctional interactions between denaturant molecules and protein groups can serve to cross-link different parts of the latter through noncovalent interactions so as to stiffen them.[24,25] Contraction of proteins at subdenaturing amounts of chemical dentaturants can be inferred from appropriate functional assays or by direct relaxation measurements. For example, shrinkage of hydrodynamic volumes of lysozyme and BSA at subdenaturing GdnHCl has been observed by ^{1}H NMR relaxation dispersion measurements.[26] More recent studies confirm that hydrodynamic radius of lysozyme at pH 5 and low urea decreases.[27] Decrease in the apparent thermal expansion coefficient of subdenatured staphylococcal nuclease at low urea has also been reported.[28] Considering these reports, we suspect from the results of urea dependence of $<\alpha>$ that pretranstion state of *At*PP1 is slightly more contracted than the native state. Strengthening of hydrogen bonds, and hence smaller values of $<\alpha>$ in the presence of urea, is a manifest of denaturant-induced protein contraction. It is important to note however that the binding effect of urea on the stiffness and motions of different parts of the protein is highly irregular. Hence, expansion of hydrogen bonds under a given condition of the denaturant is expected to vary from one segment to another. The denaturant effect presented here by the mean expansion coefficient ($<\alpha>$) is marginal because the urea increment is rather small. At 0.4 M urea, for example, just about one or two urea molecules will bind to the protein,[29] which is not expected to produce a large degree of intraprotein cross links. To summarize, suppression of hydrogen bond thermal expansion coefficient over 283 to 313 K at subdenaturing levels of urea arises from denaturant-imposed restraints on segmental motions of the protein.

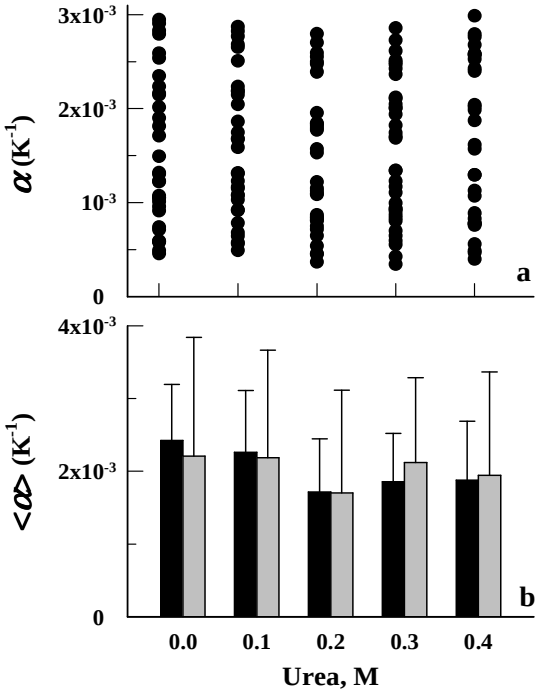


Figure 7. Variation of thermal expansion coefficient at subdenaturing concentrations of urea, pH 3.5. (a) The total spread in values of expansion coefficient of all measured hydrogen bonds as a function of urea concentration. (b) The mean of the spread, <α>, for β-sheet hydrogen bonds (black bars) and the protein average hydrogen bonds (grey bars). Error bars represent standard deviations.

4.4.5 Residues of *At*PP1 Do Not Access Alternative Conformational States Under Subdenaturing Conditions. Since the $^{1}H^{N}$ chemical shifts are studied with temperature and urea variables, the shift results should also be considered in terms of perturbation of secondary and tertiary structures by temperature and urea variables (Figure 1). At 313 K, which is the highest temperature used here, ~50% of the native-state (0 M Urea) tertiary structure is lost, and the loss is greater as the urea content is raised up to 0.4 M. Increment of urea from 0 to 0.4 M also perturbs the secondary structure to a small extent (Figure 5). In spite of these perturbations, δ_H for all

resonances vary linearly rather than curving out (Figure 3), indicating no influence of tertiary structure content on the hydrogen bond geometry. The results also suggest that any structural readjustment with varying solvent conditions do not produce locally unfolded alternative conformations, or if produced are not accessible to the amide sites. Accessibility and population of such high-energy alternative states are expected to produce curvatures in the temperature-δ_H graphs, a proposal due to Williamson and coworkers.[7,15,16] Apparently, the temperature and denaturant conditions employed to measure the chemical shifts are rather mild when compared with the denaturant concentration required for global unfolding characterized by secondary structure dissolution (Figure 1), which do not allow for population of alternative conformational states.

4.5 Summary and Conclusion

Temperature increment in the 283-313 K range produces linear downfield shift of hydrogen-bonded $^1H^N$ resonances irrespective of their location in different types and segments of secondary structure. This is atypical of β-sheet amide protons whose resonances generally move upfield due to weakening of hydrogen bonds at high temperature. The reason for this behavior is not quite understood, but could rest on the details of hydrogen bond geometry. In particular, thermally driven transverse vibrations could change the H···O=C angle (θ_2) affecting the direction of oxygen lone pair orbitals. The absence of detectable curvature in chemical shift-temperature graphs may be due to non-accessibility of the amide sites to alternative conformational states. The protein-average value of linear thermal expansion coefficient over 283 to 313 K is 2.204×10^{-3} K^{-1}, which is somewhat larger than

observed for other proteins under matching temperature conditions. This is possibly due to intrinsically disordered character of *At*PP1. The detectable decrease in average thermal expansion coefficient of *At*PP1 upon incrementing urea from 0 to 0.4 M urea supports the general idea that denaturant binding stiffens segments of the protein to an extent depending on the location of binding sites and the number of urea molecules bound.

4.6 References

1. Jeffrey, G. A;Saenger, W. *HydrogenBondinginBiologicalStructures*; Springer-Verlag: Berlin, Germany, 1991.
2. Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press, Oxford, U.K., 1999.
3. Grzesiek, S.; Cordier, F; Jaravine, V.; Barfield, M.Insights into Biomolecualr Hydrogen Bonds from Hydrogen Bond Scalar Couplings. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *45*, 275-300.
4. Ohnishi, M.; Urry, D. W.Temperature Dependence of Amide Proton Chemical Shifts: The Secondary Structures of Gramicidin S and Valionomcin. *Biochem. Biophys. Res. Commun.***1969**, *36*, 194-202.
5. Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. Folding of Immunogenic Peptide Fragments of Proteins in Water Solution. I. Sequence Requirements for the Formation of a Reverse Turn. *J. Mol. Biol.* **1988**, *201*, 161-200.
6. Skalicky, J. J.; Selsted, M. E.; Pardi, A. Structure and Dynamics of the Neutrophil Defensins NP-2, NP-5, and HNP-1: NMR Studies of Amide Hydrogen Exchange Kinetics. *Proteins***1994**, *20*, 52-67.
7. Baxter, N. J.; Williamson, M. P. Temperature Dependence of ^{1}H Chemical Shifts in Proteins. *J. Biomol. NMR***1997**, *9*, 359-369.
8. Cordier, F.; Grzesiek, S. Direct Observation of Hydrogen Bonds in Proteins by Interresidue $^{3h}J_{NC'}$ Scalar Couplings. *J. Am. Chem. Soc.* **1999**, *121*, 1601-1602.
9. Cornilescu, G.; Hu, J-S.; Bax, A. Identification of Hydrogen Bonding Network in a Protein by Scalar Couplings. *J. Am. Chem. Soc.* **1999**, *121*, 2949-2950.

10. Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A. Correlation Between $^{3h}J_{NC'}$ and hydrogen Bond Length in Proteins. *J. Am. Chem. Soc.* **1999**, *121*, 6275-6279.
11. Cordier, F.; Grzesiek, S. Temperature-dependence of Protein Hydrogen Bond Properties as Studied by High Resolution NMR. *J. Mol. Biol.* **2002**, *715*, 739-752.
12. Hong, J.; Jing, Q.; Lishan, Y. The Protein Amide $^{1}H^{N}$ Chemical Shift Temperature Coefficient Reflects Thermal Expansion of the N−H···O=Cof Hydrogen Bond. *J. Biomol. NMR***2013**, *55*, 71-78.
13. Wishart, D. S.; Sykes, B. D.; Richards, F. M. Relationship Between Nuclear Magnetic Resonance Chemical Shift and Protein Secondary Structure. *J. Mol. Biol.* **1991**, *222*, 311-333.
14. Xoconostle-Cazares, B.; Xiang, Y.; Ruiz-Medrano, R.; Wang, H. L.; Monzer, J.; Yoo, B. C.; McFarland, K. C.; Franceschi, V. R.; Lucas, W. J. Plant Paralog to Viral Movement Protein that Potentiates Transport of mRNA into the Phloem. *Science***1999**, *283*, 94-98.
15. Williamson, M. P. Many Residues in Cytochrome *c* Populate Alternative States Under Equilibrium Conditions. *PROTEINS*:*Struct.Funct. Genet.***2003**, *53*, 731-739.
16. Tunnicliffe, R. B.; Waby, J. L.; Williams, R. J.; Williamson, M. P. An Experimental Investigation of Conformational Fluctuations in Proteins G and L. *Structure***2005**, *13*, 1677-1684.
17. Semisotnov, G. V.; Rodionova, N. A.; Razgulayev, O. I.,; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. Study of the Molten Globule Intermediate State in Protein Folding by a Hydrophobic Fluorescent Probe. *Biopolymers***1991**, *31*, 119-128.

18. Wagner, G.; Pardi, A.; Wuthrich, K. Hydrogen Bond Length and ^{1}H NMR Chemical Shifts in Proteins. *J. Am. Chem. Soc.* **1983**, *105*, 5948-5949.
19. Del Bene, J. E.; Perera, S. A. Bartlett, R. J. Hydrogen Bond Types, Binding Energies, and ^{1}H NMR Chemical Shifts. *J.* Phys. Chem. A **1999**, *103*, 8121-8124.
20. Barfield, M. Structural Dependencies of Interesidue Scalar Coupling $^{3h}J_{NC'}$ and Donor ^{1}H Chemical Shifts in the Hydrogen Bonding Regions of Proteins. *J. Am. Chem. Soc*. **2002**, *124*, 4158-4168.
21. Parker, L. L.; Houk, A. R.; Jensen, J. H. Cooperative Hydrogen Bonding Effects are Key Determinants of Backbone Amide Proton Chemical Shifts in Proteins. *J. Am. Chem. Soc*. **2006**, *128*, 9863-9872.
22. Kjaergaard, M.; Norholm, A.; Hendus-Altenburger, R.; Pedersen, S. F.; Pulsen, F. M.; Kragelund, B. B. Temperature-dependent Structural Changes in Intrinsically Disordered Proteins: Formation of –helices or loss of Polyproline II?*ProteinSci*.**2010**, *19*, 1555-1564.
23. Sass, H. J.; Schmid, F. F.; Grzesiek, S. Correlation of Protein Structure and Dynamics to Scalar Couplings Across Hydrogen Bonds. *J. Am. Chem. Soc*. **2007**, *129*, 5898-5903.
24. Bhuyan, A. K.Protein Stabilization by Urea and Guanidine Hydrochloride. *Biochemistry***2002**, *41*, 13386-13394.
25. Kumar, R.; Prabhu, N. P.; Yadaiah, M.; Bhuyan, A. K. Protein Stiffening and Entropic Stabilization in the Subdenaturing Limit of Guanidine Hydrochloride. *Biophys. J.* **2004**, *87*, 2656-2662.
26. Rao, M. T.; Bhuyan, A. K.; Venu, K.; Sastry, V. S. Nonlinear Effect of GdnHCl on Hydratin Dynamics of Proteins: a ^{1}H Magnetic Relaxation Dispersion Study. *J.*

Phys. *Chem*. B **2012**, *113*, 6994-6702.

27. Yasin, U. M.; Sashi, P.; Bhuyan, A. K. Free-energy Landscape of Lysozyme: Multiple Near-native Conformational States and Rollover in the Urea Dependence of Folding Energy. *J*. *Phys*. *Chem*. B **1993**, *118*, 6662-6669.
28. Ravindra, R.; Royer, C.; Winter, R. Pressure Perturbation Calorimetric Studies of the Solvation and the Thermal Unfolding of Staphylococcal Nuclease. *Phys*. *Chem*. *Chem*. *Phys*. **2004**, *6*, 1952-1961.
29. Makhatadze, G. I.;Privalov, P. L. Protein Interactions with Urea and Guanidinium Hydrochloride: A Calorimetric Study.*J*. *Mol*. *Biol*. **1992**, *226*, 491-505.

Chapter-5

Expansion and Internal Friction in Unfolded Protein Chain

5.1 Abstract

Similarities in global properties of homopolymers and unfolded proteins provide approaches to mechanistic description of protein folding. Here, hydrodynamic properties and relaxation rates of the unfolded state of carbonmonoxide-liganded cytochrome *c* (cyt-CO) have been measured using NMR and laser photolysis methods. Hydrodynamic radius of the unfolded chain gradually increases as the solvent turns increasingly better, consistent with theory. Curiously, however, the rate of intrachain contact formation also increases with increasing denaturant concentration, which, by Szabo, Schulten, and Schulten theory for the rate of intramolecular contact formation in a Gaussian polymer, indicates growing intramolecular diffusion. It is argued that diminishing non-bonded atom interactions with increasing denaturant reduces internal friction, and thus increases the rate of polypeptide relaxation. Qualitative scaling of the extent of unfolding with non-bonded repulsions allows description of internal friction by a phenomenological model. The degree of non-bonded atom interactions largely determines the extent of internal friction.

5.2 Introduction

Polymer theory suggests that denatured protein chains should continue to expand gradually even after completion of the structural transition as denaturant concentration is increased further.[1] This should happen because the chain volume

increases when the solvent quality becomes increasingly better, favoring stronger chain-solvent interactions.[2] The increment of hydrated chain volume gives rise to higher intrinsic viscosity of the protein sample. Such expansion would obviously affect thermodynamic and transport properties of the unfolded chain ensemble in a denaturant dependent manner that in turn would affect dynamics of desolvation and chain contraction processes during folding. Detailed information of the influence of denaturants on unfolded chain hydrodynamics and diffusive properties are therefore necessary. This area has been illuminated during the past decade largely by single-molecule fluorescence studies.[3-9] Another related subject of emerging interest is internal friction, whose role in protein conformational changes and folding has been recognized by a relatively fewer studies.[10-15] Recent findings indicate that chain compaction from the unfolded state of a protein can reduce intramolecular diffusion coefficient hundreds of times,[13] although the coefficient need not be uniformly distributed within the molecule.[15] Thus, understanding intramolecular friction, which can originate from diverse non-bonded atomic interactions and barriers to torsional motions of bonded atoms, is necessary for a fuller description of dynamics, folding, and function of proteins.

To study the effect of increasingly better solvent on hydrated chain volume in the random coil state of a protein, and the bearing thereof on intramolecular friction, we examine here ensemble average results of NMR and pulsed laser photolysis experiments with the carbonmonoxide complex of cytochrome *c*, both as a function of unfolding concentrations of guanidinium hydrochloride (GdnHCl) up to the limit of its aqueous solubility. The NMR experiments provide information about chain expansion, quantified by the global parameter hydrodynamic radius, $<R_H>$, and the

laser photolysis data show the variation in the rate of intrachain contact formation. Cytochrome *c* was chosen for these experiments, because it not only is an exemplary protein system, but also allows for extrinsic ligand binding to the heme group which facilitates observation of chain dynamics following photolysis of the ligand.[16] The ligand photolysis approach has, in fact, been used in a number of studies pertaining to ultrafast folding and diffusive dynamics of the protein chain.[17-21] The present study shows continuous expansion of the random coil chain and growing rate of intrachain contact formation when the protein is subjected to increasingly unfolding conditions by incrementing the denaturant concentration. By considering the growth of non-bonded atomic repulsions with loss of protein structure we provide a phenomenological model of internal friction. Overriding importance of van der Waals and hydrophobic interactions in intramolecular friction is discussed.

5.3 Experimental Section

Deuterated guanidinium chloride (GdnDCl) was prepared by repeated lyophilization of D_2O solution of the denaturant. All experiments were performed in 100 mM phosphate buffer, pH 7, 22(±0.5)°C, containing ~20 mM sodium dithionite. For non-NMR experiments the dithionite concentration was reduced to ~2 mM, and GdnDCl was replaced by GdnHCl. Denaturant concentration was determined by using an Abbe-type refractometer.

5.3.1 Equilibrium Unfolding Measurements. Fluorescence-monitored equilibrium unfolding experiments involved GdnHCl titration of dithionite-reduced

cytochrome *c* under 1 atm CO pressure. The final protein and dithionite concentrations were ~10 and 2 μM, respectively. The sample vials sealed with rubber septa were incubated at ~23°C for 1 hour before fluorescence measurement. Emission spectra (excitation: 280 nm) were recorded using a FluoroMax 4P fluorimeter at 22(±0.5)°C. Normalized intensity at λ_{max} as a function of GdnHCl was fitted to a two-state folding-unfolding model.[22]

5.3.2 NMR Spectroscopy. Details of pulsed field-gradient (PFG) NMR measurements have been described previously.[23] Briefly, ~1 mM cytochrome *c* unfolded in 4.5-8.1 M range of GdnDCl was taken in the NMR tube, deaerated, and reduced under argon or nitrogen by adding ~20 mM sodium dithionite. Carbonmonoxide was gently bubbled into the reduced cytochrome *c* solution for a minute, and the tube was sealed by using rubber stopper and teflon. PFG-NMR diffusion measurements were done in a 500 MHz Bruker spectrometer. The diffusion gradient (*z*-gradient) was varied from 3 to 50 Gauss cm^{-1}. About 1 mM 1,4-dioxane was included in the unfolded protein solution as an internal standard for R_H. Values of R_H were calculated by

$$\mathrm{I(g)} = A\exp(-\mathrm{kg}^2) \quad (1)$$

$$\mathrm{R_H^{protein}} = \mathrm{R_H^{dioxane}}\left(\frac{\mathrm{k_{dioxane}}}{\mathrm{k_{protein}}}\right) \quad (2)$$

where, *I* is the NMR signal intensity, g is the gradient strength, and the decay constant, *k*, is proportional to the diffusion coefficient.

5.3.3 Flash Photolysis. Kinetics of intrachain contact formation was

measured by laser photolysis.The sample (~10 μM) contained in a quartz cuvette was liganded with CO by the same procedure as above, except that the concentration of dithionite was cut down to ~2 mM. Heme-bound CO was photolyzed by irradiating with 50 μJ pulses of the second harmonic of a Spectra-Physics 10 Hz YAG laser. To ensure photolysis each time in fresh volume of sample, the protein solution was circulated by a pump arrangement. Chain relaxation after each photolysis pulse was probed by employing a pulsed Xe lamp at a single wavelength. The basic configuration of the spectrometer is based on Applied Photophysics flash photolysis instrument. Relaxation traces were analyzed by exponential fits of time-base data.

5.4 Results and Discussion

5.4.1 Unfolding of Ferrocyt *c* in the Presence of CO. The experimental system used here was developed initially to study ultrafast protein folding. The system relies on binding of carbonmonoxide to the heme of denatured ferrocytochrome *c* so as to shift the folding-unfolding equilibrium (N⇌U) toward the unfolded state through mass action effect.[16] Briefly, CO preferentially binds to unfolded ferrocytochrome *c*, and the N⇌U transition under 1 atm CO pressure is characterized by a transition midpoint of ~3.8 M GdnDCl at ~23°C (Figure 1a). The unfolded chain of ferrocyt *c* bound with CO, called cyt-CO, is being used here to study the dimensional properties and intramolecular dynamics in the 4.4 to 8.3 M range of GdnHCl, marked by heavy line in Figure 1a.

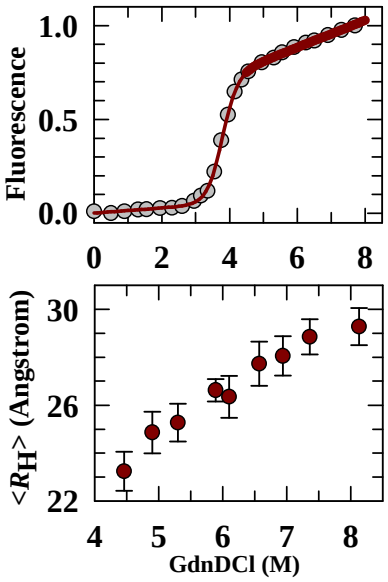
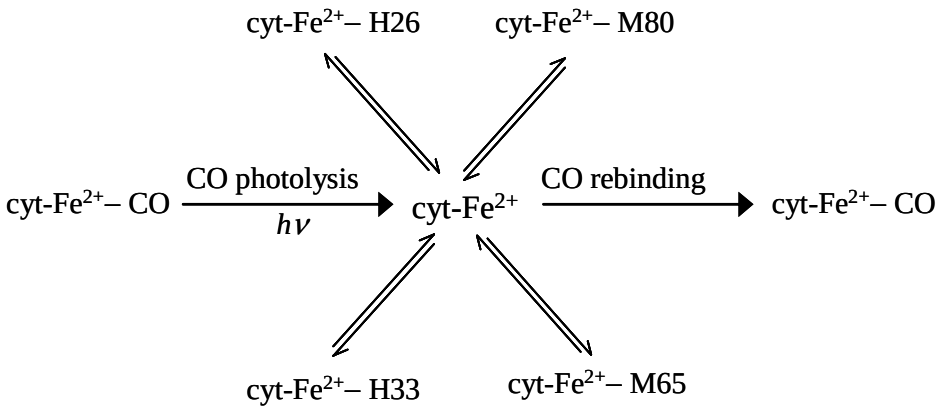


Figure 1. Continuous expansion of the unfolded chain of cyt-CO. (a) In the guanidinium chloride (GdnHCl) induced unfolding monitored by W59 fluorescence, the global transition is complete at ~4.2 M denaturant, but fluorescence continues to increase in the post-transition baseline region due to increasing distance between W59 and the heme so that fluorescence quenching by resonance energy transfer between the two decreases progressively. The solid line is the fit to a two-state model[22] yielding transition midpoint $C_m \approx 3.8$ M. (b) Variation in NMR-measured hydrodynamic radius $\langle R_H \rangle$ of the unfolded state at different concentrations of GdnDCl.

5.4.2 Expansion of Unfolded cyt-CO Chain. Global properties such as gyration radius (R_G) and end-to-end distance are widely used to describe unfolded chain relaxation dynamics and coil-globule transition in protein collapse.[24] While determination of R_G has relied largely on fluorescence[4-9,25,26] and small-angle x-ray scattering (SAXS) methods,[26-30] we used pulsed field gradient NMR (see Supporting Information Figure S1) to directly measure the hydrodynamic radius R_H, which is roughly $0.64R_G$ for a Gaussian chain in good solvent. The denaturant increment from 4.4 to 8.3 M GdnDCl increases the $\langle R_H \rangle$ of cyt-CO from 23.25 to 29.28 Å (Figure 1b), suggesting progressive weakening of interresidue interactions and hydrophobic forces between side chains. The denaturant dependent expansion of the unfolded cyt-

CO chain is consistent with polymer theory,[2] and has also been reported for some proteins by single-molecule fluorescence studies as well as by all-atom simulation.[4,7] Commendably, a hint at the denaturant dependent expansion of unfolded cyt *c* was also provided nearly 40 years ago by Tsong.[31] However, a SAXS study has shown that R_G of unfolded cyt *c* is little sensitive to GdnHCl concentration,[29] inconsistent with the present result as well as the mounting evidence for continuous chain expansion with increasing concentration of the denaturant. The reason for this discrepancy between NMR and x-ray scattering results is uncertain at present, even as a recent study finds that denaturant dependence of R_G determined by fluorescence methods is incongruent with that obtained from SAXS.[26] The results presented here nonetheless are consistent with the polymer tenet as well as the general belief that polypeptide chains should expand and contract with increasing and decreasing level of the denaturant.

5.4.3 Higher Rate of Intrachain Contact Formation with Better Quality of Solvent. How does expansion of the unfolded chain affect intramolecular diffusive dynamics? This is addressed by probing the rate of intrachain contact formation at increasing concentrations of the denaturant. The experiment involves monitoring the chain dynamics after photodissociation of the CO ligand of cyt-CO by a laser pulse. Based on earlier studies,[16,20] the scheme below depicts post-photolysis events whose record provides the rate of intramolecular contact formation between two sites of the polypeptide chain.



Following photodissociation, side chains of all four intrapolypeptide ligands can bind to the heme iron, a process that involves loop formation between H18, to which the heme is bound, and the respective residues.The post-photolysis chain relaxation processes were probed in the time bin straddling nano and microsecond regimes. Kinetics in this bin are due to binding of intrachain methionines (M65 and M80) and histidine (H26 and H33) ligands to the sixth coordination site of the heme iron made available by photolyzing the CO. Bimolecular rebinding of CO that sets in much later and extends into milliseconds does not interfere with the kinetics of chain relaxation considered here (see Supporting Information Figure S2). The trace for the 0.06-20 μs bin shows that relaxation of the unfolded chain is characterized by two exponentials (Figure 2a), the faster one (λ_1) is associated with binding of M80 and M65, and the slower one (λ_2) due to binding of H26 and H33.[19,20] These transient association-dissociation events cease to occur after CO rebinds from the solvent (Scheme 1). Figure 2b shows the dependence of λ_1 and λ_2 on unfolding concentrations of

GdnHCl. Since concentrated solutions of GdnHCl are viscous, the data presented were corrected for the effect of solvent viscosity on the relaxation rate constants (see Supporting Information for the procedure).

For an intuitive interpretation, we consider the depiction in Figure 2c based on the NMR results of denaturant-induced chain expansion, where the heme-ligand distances gradually increase as the protein is increasingly unfolded. These distances can be approximated if the mean-square end-to-end distance, $\langle r^2 \rangle$, according to Flory characteristic ratio is assumed to provide also the mean-square distance between two residues of a Gaussian segment of the polypeptide chain.[17] The distance is then given by $\langle r^2 \rangle \sim L^2 \sim C_n n l^2$, where C_n is the characteristic ratio, and $l \approx 3.8$, which is the length between successive α-carbons along the backbone chain.[32]In their original study, Brant and Flory simulated 'good solvent' conditions by using aqueous, phenolic, and acidic solvents to determine C_n values of homopolymers of aspartate, lysine, and glutamate. For the three polymers in aqueous acidic and basic solvents, they arrived at numerical values of C_n in the 8.3-8.8 range, each with ~10% error.[32] The C_n ratio is related to the intrinsic viscosity of the protein sample, $\eta_{int}=(\eta_{sp}/c)$ for the limit $c \rightarrow 0$, where the specific viscosity $\eta_{sp}=(\eta-\eta_o)/\eta_o$, η and η_o being bulk viscosity of the protein solution, and the solvent viscosity, respectively. In the range of GdnHCl concentration used in our experiments (4.4 to 8.3 M), where the protein is already in the unfolded state, the variation in η_{int} across the denaturant range may not besubstantial. Assuming that the increase in η_{int} with GdnHCl will be accommodatedwithin the uncertainty range of 10%, we chose to use $C_n \approx 8.5$, which is the average of the values for the three homopolymers in aqueous medium that Brant and Flory prescribed.[32]

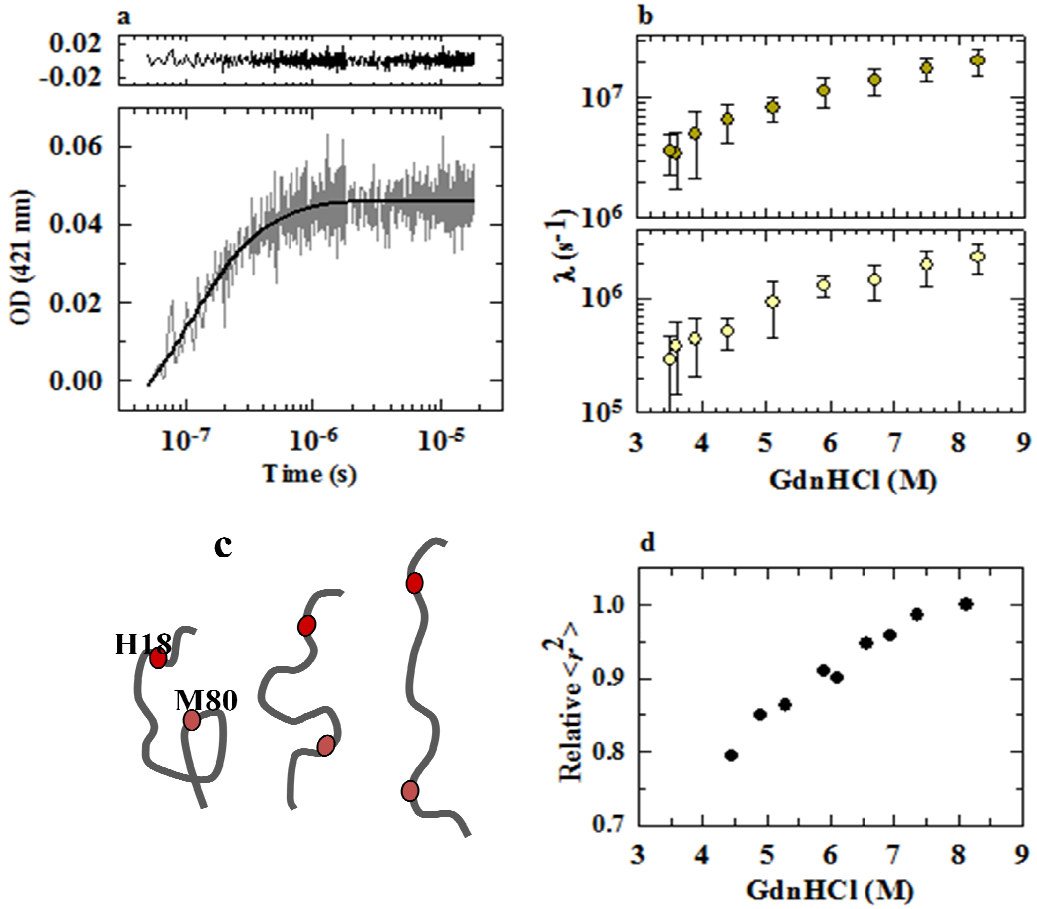


Figure 2. Relaxation kinetics and chain expansion. (a) Post-photolysis kinetics is described by two-exponentials corresponding to heme-methionine (fast) and heme-histidine (slow) contact formation. (b) The observed relaxation rate for both fast (black) and slow (blue) phases increase with increasing denaturant concentration. (c) The distance between H18, to which the heme is bonded, and M80 increases with increasing unfoldedness (from left to right). (d) Relative H18-M80 distance approximated by $\langle r^2 \rangle \sim L^2 \sim 6(R_H/0.64)^2$.

By using $C_n \approx 8.5$, the distance between H18, to which the heme moiety is bound, and M80 is estimated to ~86 Å in the random coil state. In the present context, the random coil state is somewhat arbitrary because all unfolding concentrations of the denaturant produce random coil states. For a qualitative scaling of denaturant dependence of the distance between H18 and M80 in the random coil states, we operationally use the empirical relation $\langle r^2 \rangle = 6R_G^2$, where R_G is the radius of gyration,[17] as $\langle r^2 \rangle \sim L^2 \sim 6(R_H/0.64)^2$ and obtain the denaturant dependence of the distance between H18 and M80 on a relative scale (Figure 2d). These estimates strongly suggest gradually growing distance between the heme and M80 as the denaturant level is incremented.

The present result of denaturant dependence of intrachain contact formation in unfolded cyt *c* chain is totally inconsistent with that provided in an earlier study, which suggested that the rate actually decreases as the GdnHCl concentration increases.[18] In that study, singular value decomposition (SVD) was used to reformulate sets of transient spectra, while the present work relied on single-wavelength real-time signals recorded by the use of a photomultiplier tube. Although such differences are unlikely to contribute to the discrepancy, that study argued that decreasing rate of intrachain loop formation with increasing denaturant concentration observed there was consistent with the SAXS result of Segel et al.[29] But as discussed already, SAXS reveals little dependence of R_G on denaturant level. The present results of $\langle R_H \rangle$ and intrachain contact formation rate, both growing with denaturant concentration, are consistent with each other because internal diffusion is expected to increase as the protein is incrementally unfolded.

5.4.4 Intrachain Contact Formation Rate Increases Due to Increased Internal Diffusion. Binding of an intrapolypeptide ligand, say M80, to the heme involves large-scale diffusive motion of the chain that brings the heme and the ligand to the appropriate contact distance, *a*, defined by the potential of mean force. The contacting heme and M80 can now form a bond, or the chain segments may part away by diffusion. The rate of actual chemical bond formation is generally faster than the rate of diffusion because the bond between a geminate pair of reactants can form very rapidly with a τ value of the order of a nanosecond or less, implying that it is the diffusion rate which limits the association process. Similarly, the dissociation process entails bond breaking to produce a geminate pair, which may form the bond again or diffuse away. Details of these processes can be found in earlier photolysis studies.[17,20] For the present, we simply take λ as a combination of limiting rate constants for association and dissociation ($\lambda \sim k_+ + k_-$).

The rate of contact formation between two sites of a polymer chain depends on several factors, including stiffness, chain volume and dimension, and relative diffusion of chain sites. For an ideal (Gaussian) chain lacking intrachain interactions, the rate of contact formation is obtained from the theory of Szabo, Schulten, and Schulten.[33] For a contact distance much smaller than $\langle r^2 \rangle^{1/2}$, the rate constant is given by

$$\lambda = 3\left(\frac{6}{\pi}\right)^{1/2} \frac{Da}{\langle r^2 \rangle^{3/2}} \qquad (3)$$

where the contact distance a~4 Å, and D is the end-to-end diffusion constant. Both λ_1 and λ_2 measured in our experiment increase as the protein is increasingly unfolded at higher denaturant concentration (Figure 2b), which, by the equation above, means

that the diffusive motions of the chain segments that bring the interacting sites to within the contact distance must also increase. After removing the contribution of solvent-dependent D by correcting for solvent viscosity effect, one finds a 6-fold increase in values of both λ_1 and λ_2 as GdnHCl concentration is increased from 4.4 to 8.3 M (Figure 2b). Thus, the increase of intrachain contact formation rates originates from increased internal diffusive motions of the chain as it expands in higher levels of the denaturant.

5.4.5 Chain Expansion, Non-bonded Atom Interactions and Internal Diffusion. The result above prompts one to consider the nature of D in view of conformational and configurational aspects of proteins across the folding-unfolding equilibrium. It is generally believed that protein folding and conformational changes are affected by not only solvent friction, but also internal friction.[10,11,13,15] To invoke the importance of internal friction in our problem, we assume that attractive interactions between non-bonded atoms within the polypeptide backbone as well as side chains decrease as the protein is progressively unfolded by incrementing the denaturant. This is conceivable because hydrophobic interactions are increasingly weakened and any residual structure is largely eliminated at higher denaturant concentration. The importance of van der Waals repulsion in the rotational potential about the polypeptide backbone was, in fact, realized many years ago.[34]

To provide an analytical model, we consider non-bonded atom interactions to be the major factor which scale internal diffusion. These interactions are removed as the protein unfolds, and we assume that the upper limit of non-bonded repulsions is reached when the protein is 'absolutely unfolded', i.e., when the chain density (ρ_1),

defined as the number of monomers per unit volume, approaches zero. Now, the functional dependence of internal friction on non-bonded atom attractions, $f(n)$, can be represented by the series

$$f(n) = f_0 + f_1 n + f_2 n^2 + \cdots \quad (4)$$

where, f_o is the internal friction when no non-bonded atom attraction exists ($\rho_i \to 0$), n is the extent of non-bonded attraction, and f_i are coefficients. For $n \to 0$, $f(n)=f_o$, the situation of 'absolute unfoldedness' ($\rho_i \to 0$). Protein unfolding is associated with decrease of n, while refolding proceeds with growth of n. The overall diffusion coefficient is simply

$$D = D_{\text{ext}} + \frac{k_B T}{f_0 + f_1 n + f_2 n^2 + \cdots} \quad (5)$$

where, D_{ext} is the solvent-damped external diffusion.

To apply this model of internal diffusion, one needs to have a reference where $\rho_i \to 0$, so that $f(0)=f_o$. Experimental conditions where $n \to 0$ may be difficult to obtain, because the protein chain continues to expand. For the purpose here, such ideal behavior may be assumed at best when GdnHCl concentration is 8.3 M, which is the limit of aqueous solubility of the denaturant (Figure 1). If non-bonded atom attractions grow as the denaturant concentration is lowered, by corollary, non-bonded repulsions should increase as the denaturant is increased. By normalizing the experimental concentrations of GdnHCl with reference to the highest concentration employed (8.3 M) one can set the x-scale in 0 to 1 interval. In the native state, where non-bonded interactions are all attractive, there is no repulsion (x=0 and n=1), and in the 'absolutely unfolded' state, all interactions are repulsive (x=1 and n=0). The x-dependence of normalized λ_1 with the first two coefficients, f_1 and f_2 is shown in

Figure 3. The fit indicates that the first-order coefficient f_1 is double the value for f_2. Inclusion of higher order coefficients will produce a fit approaching the origin where highly overdamped internal dynamics contribute little to the chain relaxation rate.

5.4.6 Phenomenological Nature of the Internal Diffusion Model. The premise of friction-dependent relaxation discussed here is similar to that of de Gennes who suggested that relaxation times in polymer chains have contributions from both solvent damped motions and intramolecular dynamics.[11,35,36] While intramolecular drag forces could originate from several mechanisms, including torsional potentials for bond rotations, collisions between neighboring atoms, and transient long-range interactions within the molecule,[11,15,37] the model of internal friction here considers interactions amongst non-bonded atoms in a general way based on the fact that protein unfolding removes the van der Waals and hydrophobic interactions in a denaturant dependent manner. The rationale for the emphasis on short- and long-range non-boding interactions arises from the central role these interactions play in folding and structure acquisition. The origin and nature of bond rotational potentials, on the other hand, are less understood. Recent studies indicate that Coulombic interactions between bonded atoms are the major sources of torsional potentials.[38] However, electrostatic interactions between bonded atoms are unlikely to undergo a drastic change across the folding-unfolding transition so as to produce a huge difference in internal friction. To summarize, the model here weighs internal friction with ‘concentration’ of non-bonded interactions.

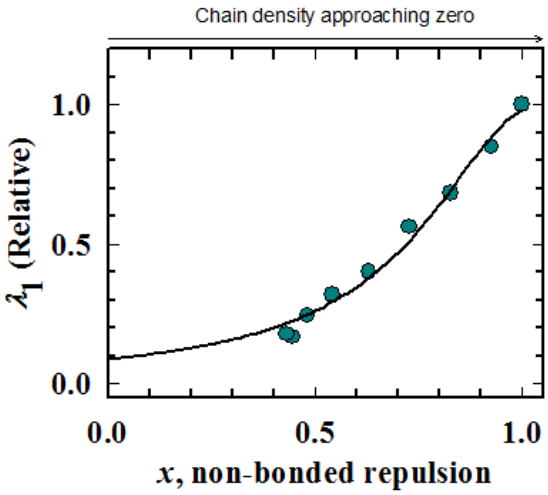


Figure 3. Non-bonded atom interaction model for internal friction. The observed rate of heme-M80 intrachain contact formation (λ_1) decreases with decreasing non-bonded repulsions (or increasing intramolecular interactions) which is the hallmark of protein folding. The denaturant concentration is along the abscissa. The highest concentration of GdnHCl achievable in this work was 8.3 M, which has been taken to produce the absolutely unfolded protein, and the density of the unfolded chain under this condition approaches zero. The solid line through experimental data is drawn as described in the text.

Attractive interactions between non-bonded atoms in native proteins can make internal friction a major factor in functional and dynamical properties of proteins. A point of particular interest would be the effect of temperature, which reduces external protein-solvent friction because solvent viscosity is highly temperature dependent. However, intramolecular van der Waals and hydrophobic interactions, both of which share similar qualitative features,[39] are strengthened within the functional limit of temperature, implying higher internal friction. It is now certain that compaction of proteins accompanied by attractive non-bonded atom interactions increases internal friction substantially, implying the existence of a definite relation with entropy. To this end, a detailed theoretical treatment of internal friction is due.

5.5 References

1. Dill, K. A. Theory for the Folding and Stability of Globular Proteins. *Biochemistry***1985**, *24*, 1501-1509.
2. Dill, K. A.; Shortle, D. Denatured States of Proteins. *Annu. Rev. Biochem.* **1991**, *60*, 795-825.
3. Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the Free Energy Surface for Protein Folding with Single Molecule Fluorescence Spectroscopy. *Nature***2002**, *419*, 743-747.
4. Sherman, E.; Haran, G. Coil-Globule Transition in the Denatured State of a Small Protein. *Proc. Natl. Acad. Sci. USA***2006**, *103*, 11539-11543.
5. Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.;, Horsley, D. A.; Seckler, R.; Bakajin, O. *etal*. Mapping Protein Collapse with Single-Molecule Fluorescence and Kinetic Synchrotron Radiation Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 105-110.
6. Huang, F.; Sato, S.; Sharpe, T. D.; Ying, L.; Fersht, A. R. Distinguishing between Cooperative and Unimodal Downhill Protein Folding. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 123-127.
7. Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. Characterizing the Unfolded States of Proteins Using Single-Molecule FRET Spectroscopy and Molecular Simulations. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 1528-1533.
8. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures. *Proc. Natl. Acad. Sci. USA***2007**, *104*, 2649-2654.

9. Ziv, G.; Haran, G. Protein Folding, Protein Collapse, and Tanford's Transfer Model: Lessons From Single Molecule FRET. *J. Am. Chem. Soc.* **2009**, *131*, 2942-2947.
10. Ansari, A.; Jones, C.M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. The Role of Solvent Viscosity in the Dynamics of Protein Conformational Changes. *Science***1992**, *256*, 1796-1798.
11. Pabit, S. A.; Roder, H.; Hagen, S. J. Internal Friction Controls the Speed of Protein Folding From a Compact Configuration. *Biochemistry***2004**, *43*, 12532-12538.
12. Hagen, S. J.; Qiu, L. L.; Pabit, S. A. Diffusional Limits to the Speed of Protein Folding: Fact or Friction? *J. Phys. Cond. Mat.* **2005**, *17*, S1503-S1514.
13. Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely Slow Intramolecular Diffusion in Unfolded Protein L. *Proc. Natl. Acad. Sci. USA***2010**, *107*, 13713-13717.
14. Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Muller-Spath, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B.Quantifying Internal Friction in Unfolded and Intrinsically Disordered Proteins with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. USA***2012**, *109*, 17800-17806.
15. Borgia, A.; Wensley, B. G.; Soranno, A.; Nettels, D.; Borgia, M. B.; Hoffmann, A.; Pfeil, S. H.; Lipman, E. A.; Clarke, J.; Schuler, B.Localizing Internal Friction along the Reaction Coordinate of Protein Folding by Combining Ensemble and Single-molecule Fluorescence Spectroscopy. *Nat. Commun.* **2012**, *3*, 1195.
16. Jones, C. M.; Henry, E. R.; Hu, Y.; Chan, C. K.; Luck, S. D.; Bhuyan, A.; Roder, H.; Eaton, W. A. Fast Events in Protein Folding Initiated by Nanosecond Laser Photolysis. *Proc. Natl. Acad. Sci. USA***1993**, *90*, 11860-11854.

17. Hagen, S. J.; Hofrichter, J.; Eaton, W. A. Rate of Intrachain Diffusion of Unfolded Cytochrome *c*. *J*. *Phys*. *Chem*. *B***1997**, *101*, 2352-2365.
18. Hagen, S. J.; Carswell, C. W.; Sjolander, E. M. Rate of Intrachain Contact Formation in an Unfolded Protein: Temperature and Denaturant Effects. *J*. *Mol*. *Biol*. **2001**, *305*, 1161-1171.
19. Hagen, S.J.; Latypov, R. F.; Dolgikh, D. A.; Roder, H. Rapid Intrachain Binding of Histidine-26 and Histidine-33 to Heme in Unfolded Ferrocytochrome *c*. *Biochemistry***2002**, *41*, 1372-1380.
20. Kumar, R.; Prabhu, N. P.; Bhuyan, A. K. Ultrafast Events in the Folding of Ferrocytochrome *c*. *Biochemistry***2005**, *44*, 9359-9367.
21. Kumar, R.; Bhuyan, A. K. Viscosity Scaling for the Glassy Phase of Protein Folding. *J*. *Phys*. *Chem*. *B***2008**, *112*, 12549-12554.
22. Santoro, M. M.; Bolen, D. W. Unfolding Free Energy Changes Determined by the Linear Extrapolation Method. Unfolding of Phenylmethanesulfonyl alpha-Chymotrypsin Using Different Denaturants. *Biochemistry***1988**, *27*, 8063-8068.
23. Bhuyan, A. K. Off-pathway Status for the Alkali Molten Globule of Horse Ferricytochrome *c*. *Biochemistry***2010**, *49*, 7764-7773.
24. England, J. L.; Haran, G. Role of Solvation Effects in Protein Denaturation: From Thermodynamics to Single Molecules and Back. *Annu*. *Rev*. *Phys*. *Chem*. **2011**, *62*, 257-277.
25. Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. Distribution of End-to-End Distances of Oligopeptides in Solution as Determined by Energy Transfer. *Proc*. *Natl*. *Acad*. *Sci*. *USA***1975**, *72*, 1807-1811.
26. Yoo, T. Y.; Meisburger, S. P.; Hinshaw, J.; Pollack, L.; Haran, G.; Sosnick, T. R.;

Plaxco, K. Small-angle X-ray Scattering and Single-Molecule FRET Spectroscopy Produce Highly Divergent Views of the Low-denaturant Unfolded State. *J. Mol. Biol*. **2012**, *418*, 226-236.

27. Kataoka, M.; Hagihara, Y.; Mihara, K.; Goto, Y. Molten Globule of Cytochrome *c* Studied by Small Angle X-ray Scattering. *J. Mol. Biol*. **1993**, *229*, 591-596.
28. Kamatari, Y. O.; Konno, T.; Kataoka, M.; Akasaka, K. 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*. **1996**, *259*, 512-523.
29. Segel, D. J.; Fink, A. L.; Hodgson, K. O.; Doniach, S. Protein Denaturation: a Small Angle X-ray Scattering Study of the Ensemble of Unfolded States of Cytochrome *c*. *Biochemistry***1998**, *37*, 12443-12451.
30. Plaxco, K. W.; Millett, I. S.; Segel, D. J.; Doniach, S.; Baker, D. Chain Collapse can Occur Concomitantly With the Rate-Limiting Step in Protein Folding. *Nat. Struct. Biol*. **1999**, *6*, 554-556.
31. Tsong, T. Y. The Trp-59 Fluorescence of Ferricytochrome *c* as a Sensitive Measure of the Over-all Protein Conformation. *J. Biol. Chem*. **1974**, *249*, 1988-1990.
32. Brant, D. A.; Flory, P. J. The Configuration of Random Polypeptide Chains. I. Experimental Results. *J. Am. Chem. Soc*. **1965**, *87*, 2788-2791.
33. Szabo, A.; Schulten, K.; Schulten, Z. First Passage Time Approach to Diffusion Controlled Reactions. *J. Chem. Phys*. **1980**, *72*, 4350-4357.
34. Brant, D. A.; Flory, P. J. The Configuration of Random Polypeptide Chains. II. Theory. *J. Am. Chem. Soc*. **1965**, *87*, 2791-2800.

35. de Gennes, P. G. *ScalingConceptsinPolymerPhysics*; Cornell University Press: Ithaca, New York, 1979.
36. Portman, J. J.; Takada, S.; Wolynes, P. G. Microscopic Theory of Protein Folding Rates. II. Local Reaction Coordinates and Chain Dynamics. *J. Chem. Phys.* **2001**, *114*, 5082-5096.
37. Schulz, J. C.; Schmidt, L.; Best, R. B.; Dzubiella, J.; Netz, R. R. Peptide Chain Dynamics in Light and Heavy Water: Zooming in on Internal Friction. *J. Am. Chem. Soc.* **2012**, *134*, 6273-6279.
38. Liu, S. Origin and Nature of Bond Rotation Barriers: a Unified View. *J. Phys. Chem. A***2013**, *117*, 962-965.
39. Parsegian, V. A.; Ninham, B. W. Temperature-Dependent van der Waals Forces. *Biohys. J.* **1970**, *10*, 664-674.
40. Kawahara, K.; Tanford, C. Viscosity and density of aqueous solutions of urea and guanidine hydrochloride. *J. Biol. Chem.* 1966, *241*, 3228-3232.

5.6 Supporting Information

Correction of relaxation rates for viscosity effect of GdnHCl. The following empirical equation was obtained by fitting the viscosity data of GdnHCl solutions published by Kawahara and Tanford,[40]

$$\frac{\eta}{\eta_o} = 1 + \left(0.005x^{2.6144}\right) + \left(0.018x^{0.6594}\right) + \left(0.01213x^{0.6636}\right)$$

where η and η_o are viscosities in the presence and absence, respectively, of the denaturant concentration x. The observed chain relaxation rates were then corrected by

$$\log \lambda_{corr} = \log \lambda + \log (\eta/\eta_o)$$

Stack plots of PFG-NMR spectra

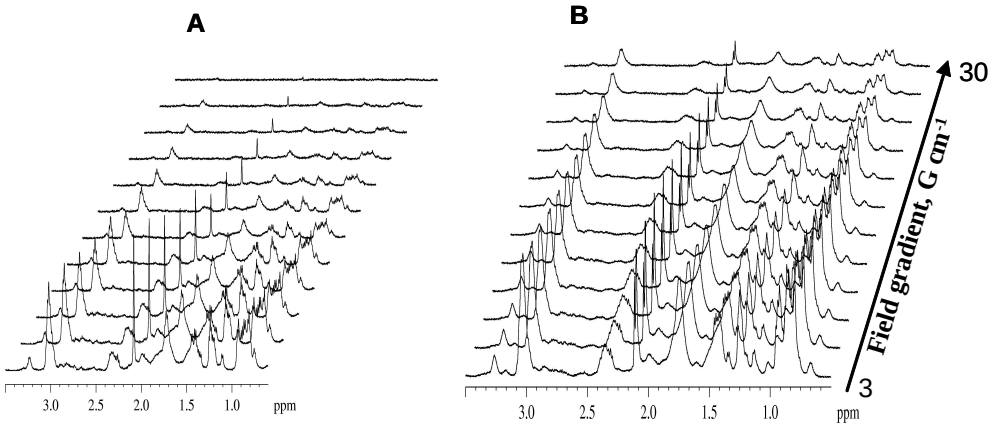


Figure S1. Aliphatic regions of PFG-sLED spectra of CO-liganded cytochrome *c* (cyt-CO) unfolded in 4.5 M GdnDCl (a) and 8.3 M GdnDCl (b). The magnetic field gradient strength (Gauss per cm) increases with the stack height.

Time bins for chain relaxation and bimolecular rebinding of CO

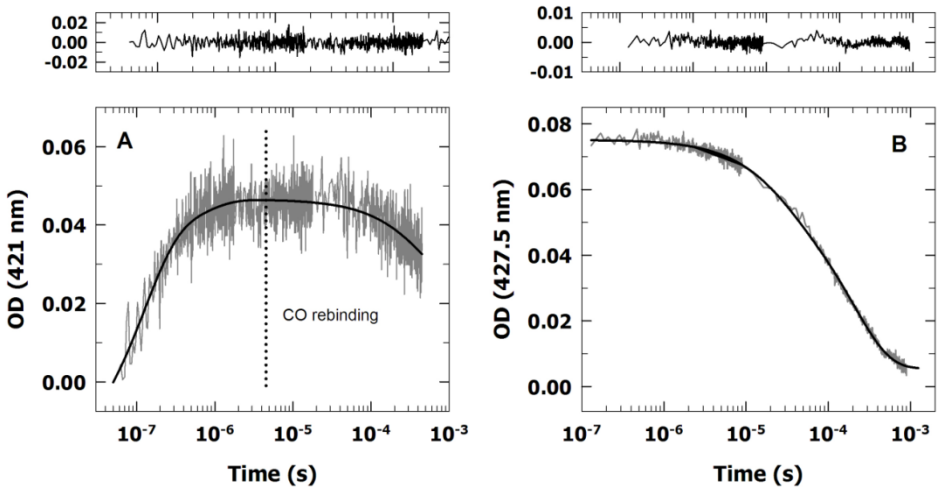


Figure S2. Post-photolysis relaxation kinetics. (A) Kinetics probed at 421 nm are described by three-exponentials, of which the slowest one (incomplete here) is assigned to rebinding of CO. The vertical dotted line provides rough indication where CO rebinding sets in. (B) The CO rebinding kinetics is better probed at 427.5 nm under the experimental conditions employed (8 M GdnHCl).

List of Publications

1. **Yasin, U. M.**; Sashi, P.; Bhuyan, A. K. Disorder-to-Order Collapse of the Intrinsically Disordered Protein *At*PP1 under Subdenaturing Conditions of Urea: A ^{15}N NMR Relaxation Study. (*To be communicated*)
2. **Yasin, U. M.**; Sashi, P.; Bhuyan, A. K.Thermal Effect on Amide Proton Chemical Shifts and Linear Expansion Coefficient of Hydrogen-Bonds in an Intrinsically Disordered Plant Protein, *At*PP1.(*To be communicated*)
3. **Yasin, U. M.**; Sashi, P.; Bhuyan, A. K. Free energy landscape of lysozyme: multiple near-native conformational states and rollover in the urea dependence of folding energy.*J. Phys. Chem. B***2014,** 118, 6662-6669.
4. Sashi P, **Yasin UM**, Balasubramanian H, Sree MU, Ramakrishna D, Bhuyan AK. Preferential water exclusion in protein unfolding.*J. Phys. Chem. B* **2014,** 118, 717-23.
5. **Yasin, U. M.**; Sashi, P.; Bhuyan, A. K. Expansion and Internal Friction in Unfolded Protein Chain. *J. Phys. Chem. B***2013**, *117*, 12059-12064.
6. Sudhamalla B, Kumar M, Kumar RS, Sashi P, **Yasin UM**, Ramakrishna D, Rao PN, Bhuyan AK.Enzyme dimension of the ribosomal protein S4 across plant and animal kingdoms.*Biochim Biophys Acta.* 2013, 1830, 5335-41.
7. Sashi P, **Yasin UM**, Bhuyan AK.Unfolding action of alcohols on a highly negatively charged state of cytochrome *c*.*Biochemistry.* 2012, 51, 3273-83.
8. Yadaiah M, Rao PN, Sudhamalla B, Ramakrishna D, **Yasin UM**, Bhuyan AK.Cloning, Escherichia coli expression, purification, characterization, and enzyme assay of the ribosomal protein S4 from wheat seedlings (Triticum vulgare).*Protein Expr Purif.* 2012, 81, 55-62.