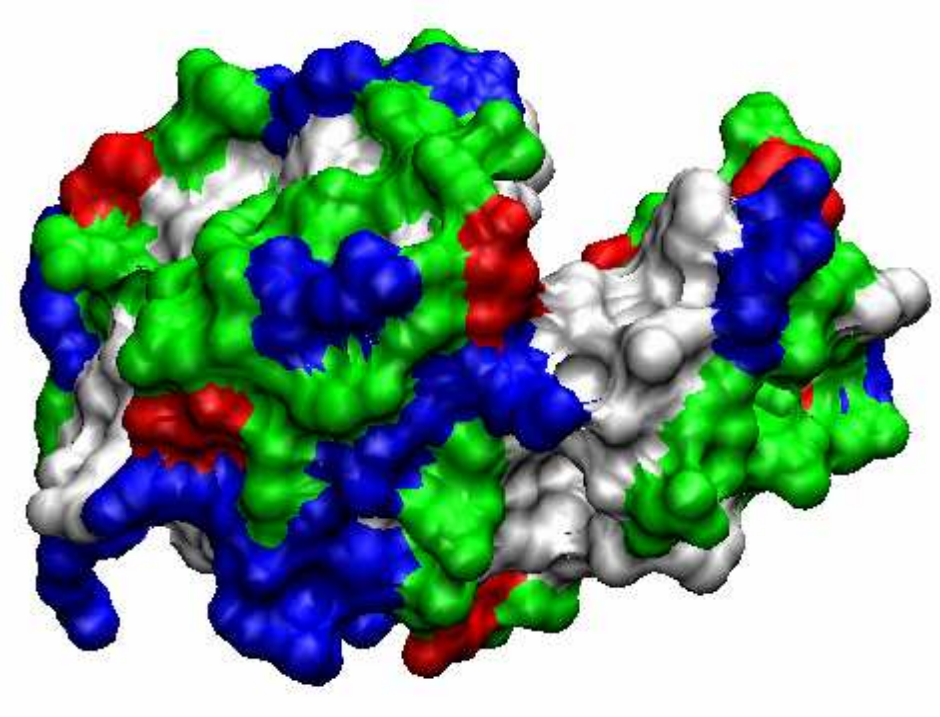
REFOLDING KINETICS OF LYSOZYME: NUCLEAR MAGNETIC RESONANCE & MOLECULAR DYNAMICS STUDY OF HEN EGG-WHITE LYSOZYME



by
ÇETİN BALOĞLU

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of the requirements for the degree of
Master of Science
Sabancı University
Spring 2005

Hen Egg-White Lysozyme, blue shows basic residues, red acidic, green polar, white non-polar

ACKNOWLEDGMENT

First of all, I would like to thank my thesis supervisor Canan Baysal for providing guidance and insight during this thesis study. She introduced me to this exiting field and clearly stated my needs. She always led me kindly when I was lost in documentation. Studying with her was a great experience to realize how a scientist should work. I believe I have much to learn from her.

I feel very happy to have known and worked with my thesis co-supervisor Alpay Taralp. He dedicated a lot of time to help me in every subject and continuously supported, motivated and encouraged me to do better. I will never forget his great support during my whole 3 years at Sabanci University.

I am very thankful to Ugur Sezerman for supporting my entry into Sabanci University during troubled times. My life has transformed for the better at Sabanci University; I will never forget his positive effect in my life.

During this thesis, I felt very lucky to meet Michel Koch. I thank him for his comments and discussions during this study. He kindly answered all my questions during my studies.

I thank Ali Rana Atılgan for his motivative support before and after my thesis presentation. His comments were so helpful.

Zehra Sayers was always kind and helpful to me. I am very happy to have known her. She also reviewed my work and had very important comments to share.

Yücel Saygın was very kind to attend my thesis defence jury. I attended his Data Mining course and learned very important methods that I still use.

I thank to Mutlu Doğruel for his friendship, scientific discussions and being a travel mate to me.

We spent 3 full years with Burcu Kaplan within the Faculty and Mutlukent. It's not possible to forget the days we worked, quarrelled, discussed, play and walked around.

We shared many secrets with my dear home mate Serkan İsmail Göktuna. He was always kind to me and helped a lot in scientific issues during this study.

Günseli Bayram Akçapınar was always ready to help and support me with her great friendship. I am so happy to have found such a nice friend.

The IT department of Sabanci University financially supported me for the last 3 years and kindly provided technical help whenever I requested. I admired their working discipline at the university. They also encouraged me much towards completing my academic responsibilities.

The existence of my lovely dad and mom, my sisters Serpil and Seda and brothers Cemil, Metin and Mehmet always gave me peace and support throughout my life. Without them, I would not have come to this point.

"If we knew what it was we were doing, it would not be called research, would it? "
Albert Einstein

ABSTRACT

NMR spectral analyses and molecular dynamics (MD) simulations in the temperature range of 300-355K and at 500K were used to probe the unfolding, refolding and gelation of native hen egg-white lysozyme. In the first part, preliminary ^{1}H experiments were conducted on samples that had been freshly dissolved in 10%D_2O / 90%H_2O and spiked with acetonitrile as reference. The samples were encouraged to unfold and refold by ramping the temperature of the NMR probe, yielding real-time spectra that reflected changes of structure. NMR experiments were also performed on lysozyme that had been labeled along the surface with isotopically enriched ^{13}C-methyl groups. This strategy, which promoted carbon observation at sensitivities approaching that of many homonuclear ^{1}H experiments, facilitated the investigation of heteronuclear Nuclear Overhauser effects, spin-lattice (T_1) and spin-spin (T_2) ^{13}C relaxation times in the modified protein. Furthermore, the thermal gelation of egg-white lysozyme was monitored in solution containing different amounts of vitamin B_1 as model incipient. Protein incubated with moderate amounts of vitamin B_1, an established enzyme cofactor, showed higher tolerance to the denaturing effects of elevated temperature. In comparison, MD simulations were used to characterize the global changes of protein structure upon thermal treatment and served as a knowledge base to carry out subsequent NMR backbone dynamics studies. The backbone dynamics of lysozyme were assessed using 2ns MD simulations between 300K 355K and 4ns for 500K, in which Cα fluctuation vector, relaxation times, heat capacity, accessible surface area and solution scattering functions were compared against related experimental findings.

In general, the results supported previous interpretations that a protein fingerprint exists, which make out distinct intermediates forming along the unfolding pathways.

ÖZET

300-355 K dereceleri içerisinde gerçekleştirilen NMR grafik analizi ve moleküler dinamik (MD) simülasyonları , tavuk yumurtası beyazı lizozomu proteininin katlanma, bozulma ve jelleşmesini inceleme amaçlı kullanıldı. İlk olarak H_1 deneyleri referans amaçlı 5µl asetonitril eklenen 10%D_2O / 90%H_2O çözeltisi ile hazırlanmış protein örneklerinde gerçekleştirildi. Örnekler NMR algılayıcısı içerisine yerleştirilen örneğin sıcaklığının artırılıp azaltılması ile bozulum ve katlanmaları sağlanarak, gerçek zamanlı yapısal değişikliklere işaret eden tek boyutlu grafikler elde edildi. NMR deneyleri aynı zamanda yüzeyi ^{13}C-metil grupları ile zenginleştirilmiş lizozom ile de gerçekleştirildi. Birçok homonükleer ^{1}H deneylerindeki hassasiyetle karbon gözlenmesini sağlayan bu yöntem, değiştirilmiş proteinin Nükleer Overhauser etkisini, spin-çevre ve spin-spin ^{13}C salınım zamanlarını gözlenmesini kolaylaştırdı. Bunun yanında lizozomun ısısal kaynaklı jelleşmesi farklı miktarlardaki B_1 vitamini ortamı model alınarak incelendi. Bir enzim kofaktörü olan vitamin B_1'le etkileşime giren protein, yükselen sıcaklığa karşı yüksek bir direnç kazandı. Protein yapısındaki, sıcaklık nedeniyle oluşan yapısal değişiklikleri karakterize etmek ve protein ile yapılan çalışmaların yönünü belirlemek amaçlı MD çalışmaları gerçekleştirildi. Lizozomun çatısal yapısı, ısı kapasitesi, suyla temas eden yüzey analizi ve çözelti dağılım fonksiyonları 300-355K derecelerinde 2ns, 500K derecesinde 4ns süren simülasyonlarla incelendi. Isı kapasitesi ile elde edilen sonuçların önceki çalışmalarla uyum içinde olduğu; NMR deneyleri ile de gösterilen, bozuluma direnç noktaları ve bozulum derecelerini gösterdiği görüldü.

Genel olarak elde ettiğimiz sonuçlardan proteinlerin bozulum aşamasında ara yapılardan kaynaklanan ve deneysel yollarla beraber moleküler dinamik simülasyonları ile de görülebilecek özel bir iz taşıdığı sonucuna varıldı.

TABLE OF FIGURES

Figure 2.2: Representation of intermolecular potential forces. 8
Figure 3.1: Molecular dynamics flowchart with NAMD code. 19
Figure 3.2: Hen egg-white lysozyme (6lyz). 20
Figure 3.3: Heat capacity versus temperature. 22
Figure 3.4: DSC trace of protein melting 23
Figure 3.5: Radius of gyration is calculated at each simulated temperature 24
Figure 3.6: Radius of gyration curve at 500K. 24
Figure 3.7: Polar surface area of hen egg-white lysozyme at 500K.. 25
Figure 3.8: Solution scattering graphs of average structures 26
Figure 3.9: Thermal fluctuations. 27
Figure 3.10: Experimental and computational fluctuations 28
Figure 3.11: Residue-by-residue fluctuations of Cα atoms of HEW lysozyme at 500 K. 28
Figure 3.12: Relaxation function of hen egg-white lysozyme 29
Figure 3.13: Stretched-exponential fit of relaxation data 30
Figure 3.14: Graphical representations of most active residues 30
Figure 3.15: Graphical representations of most active residues 301
Figure 4.1: Lysozyme ^{1}H spectrum without water suppression 35
Figure 4.2: Lysozyme ^{1}H spectrum with water suppression 36
Figure 4.3: pH effect on protein structure at 35 ˚C 37
Figure 4.4: pH effect on protein structure at 65 ˚C. 37
Figure 4.5: pH effect on protein structure at 80 ˚C. 38
Figure 4.6c: ^{1}H spectrum of lysozyme at pH 9 for 80 ˚C 39
Figure 4.9: Thiamine effect on lysozyme stability at 65˚C pH 9. 41
Figure 4.10: Lysozyme at 80˚C, pH 9 41
Figure 4.11: Heat treatment of lysozyme solution. 41
Figure 4.12: NMR tubes after heat treatment 42
Figure 4.13: Natural abundance ^{13}C spectrum, at room temperature. 42
Figure 4.14: Natural abundance lysozyme ^{13}C spectra 43
Figure 4.15: ^{13}C methylated lysozyme 44
Figure 4.16: ^{13}C methylated lysozyme 44
Figure 4.17: T_1 graph of ^{13}C relaxation of lysozyme at 25˚C 45
Figure 4.18: T_1 graph of ^{13}C relaxation of lysozyme at 35˚C. 45
Figure 4.19: T_1 graph of ^{13}C relaxation of lysozyme at 65˚C. 46

TABLE OF CONTENTS

ACKNOWLEDGMENT ..... **III**
ABSTRACT ..... **IV**
ÖZET ..... **V**
TABLE OF FIGURES ..... **VII**
TABLE OF CONTENTS ..... **VIII**
LIST OF TABLES ..... **X**
LIST OF SYMBOLS ..... **XI**
LIST OF ABBREVIATIONS ..... **XII**
1. INTRODUCTION **1**
2. THEORY **4**
2.1 Lysozyme 4
2.2 Molecular Dynamics 5
2.2.1 MD Theory 7
2.2.2 Heat Capacity 10
2.2.3 Fluctuation Vector 11
2.2.4 Relaxation Phenomena 11
2.3 Nuclear Magnetic Resonance 12
2.3.1 Overview 12
2.3.2 Magnetic Properties of Nuclei 14
2.3.3 The Chemical Shift 15
2.3.4 Spin-Lattice & Spin-Spin Relaxation 16
2.3.5 Nuclear Overhauser Effect 16
2.4 Solution Scattering 17
3. MD STUDY OF HEW LYSOZYME **18**
3.1 Overview 18
3.2 Computational Systems and Computer programs 18
3.3 System Preparation 20
3.4 Simulation Details 21
3.5 Results 21
3.5.1 Heat Capacity 22
3.5.2 Radius of Gyration 23
3.5.3 Solution Scattering Functions 26
3.5.4 B-factors 27
3.5.5 Characterizing the Heterogeneous Dynamics 29
3.5.6 Unfolding Pathways 30

4. NMR EXPERIMENTS OF HEW LYSOZYME 32
4.1 Materials 33
4.2 Conditions 33
4.2.1 Solvent System 34
4.2.2 Temperature 34
4.2.3 pH Value 34
4.2.4 Buffer 34
4.3 Results 35
4.3.1 ^{1}H Spectra Analysis at Different Temperatures 35
4.3.2 Kinetic Experiments Under Different pH Values 36
4.3.2.1 Acidic 36
4.3.2.2 Basic 38
4.4 Thiamine Effect on Protein Stability 40
4.5 Natural Abundance ^{13}C Experiments 42
4.6 ^{13}C Methylation Experiments 43
4.7 Relaxation Experiments 44
5. CONCLUSIONS & FUTURE WORK 47
6. REFERENCES 50

LIST OF TABLES

Table 2.1: The characteristic times for common molecular events 9

LIST OF SYMBOLS

a_i	Acceleration of the *i*th particle
Å	Angstrom
C	Carbon
$C(t,T)$	Relaxation function of time and temperature
C_v	Constant-volume heat capacity
D	Deuterium
F*i*	Force exerted on the *i*th particle
H	Hydrogen
$I_{(s)}$	Scattered intensity
K	Kelvin
k_B	Boltmann constant
m	Mass
N	Number of atoms
n	Number of bonds
ppm	Parts per million
R_g	Radius of gyration
r_i	Position vector of the *i*th atom
R(t)	Trajectory of interest
s	Momentum transfer
T	Temperature
T_1	Spin-lattice relaxation times
T_2	Spin-spin relaxation times
V	Total potential energy
V	Volume
β	Stretch-exponent of Kohlrausch-Williams-Watts function
λ	Wavelength
δ	Chemical shift value
Δt	Time step size
v	Frequency

LIST OF ABBREVIATIONS

BPTI	Bovine Pancreatic Trypsin Inhibitor
CPU	Central Processing Unit
CSI	Chemical Shift Index
DSC	Differential Scanning Calorimetry
FID	Free Induction Decays
HEW	Hen Egg White
MD	Molecular Dynamics
MHz	One million periods per second
MPI	Message Passing Interface
NAMD	Not Another Molecular Dynamics
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NVT	Canonical Ensemble
PBC	Periodic Boundary Condition
PDB	Protein Data Bank File
PME	Particle-Mesh Ewald
PME	Particle-Mesh Ewald
PSF	Protein Structure File
SAXS	Small Angle X-ray Scattering
SD	Steepest-Descend
SE	Schrödinger Equation
SSH	Secure shell
VMD	Visual Molecular Dynamics

1. INTRODUCTION

Protein folding continues to be one of the most exciting and immensely studied subjects of the life sciences, but a detailed understanding of mechanical and dynamical aspects is still limited. A major factor impeding the convenient investigation of this problem is the speed at which proteins can fold. *In vivo* protein folding is all but inaccessible to a detailed kinetic and mechanistic analysis. Because of the obvious complexity of the cell, many scientists have chosen to conduct *in vitro* studies, especially in light of successful denaturation-renaturation studies conducted by Linus Pauling and his colleagues [Pauling *et al.*, 1951]. In 1963, Levinthal proposed his famous paradox, arguing that if proteins assume all the possible conformational shapes during folding, they would never fold in the observed millisecond to second time scales [Levinthal, 1968]. After that, researchers realized that a protein does not assume all conformational possibilities in reaching its native state. In particular they emphasized on the directing effect of folding pathways and intermediates [Wildegger and Kiefhaber, 1997; Clark and Waltho, 1997].

Even today a precise method to probe detailed structural changes is lacking. Moreover, little is known about the sampled conformations and energetics of protein folding. Most investigations related to folding kinetics, intermediates and pathways are being realized through specialized experimental and theoretical methods. For instance, molecular dynamics (MD) simulations [Williams *et al.*, 1997; Mark and van Gunsteren, 1992; Kazmirski and Daggett, 1998] of the unfolding process, and folding simulations using simplified models [Shakhnovich, 1997] have proven to be effective probes, providing useful insight. Significant progress has been achieved on expanding the knowledge base of lysozyme using these methods.

An increasing number of experimental and theoretical studies has focused on the problem of protein dynamics, particularly with the advent of Nuclear Magnetic Resonance (NMR) Spectroscopy and renewed algorithms. NMR methods applied to protein folding have provided information about the structure and dynamics of folding intermediates. *In silico* methods, with recently developed parallel computational techniques and improved forcefields, has helped predict conformational changes and related kinetics of protein intermediates [Kuwajima and Schmid, 1984]. Indeed, experiments often do not provide the molecular level of information that is available from simulations and simulations also have inherent weaknesses. Therefore, theoreticians and experimentalist share a motive to combine their efforts, affording synergy that leads to new insights of protein folding [Sen, 2004].

In this study, protein folding dynamics has been explored using NMR and recently implemented computational techniques. Lysozyme, the first enzyme discovered, was chosen as the model protein as it contains all 20 of the usual amino acids [Canfield, 1963] and as it has already served as a model system in protein chemistry & MD studies, thus facilitating the current investigation by a virtue that a convenient and extensive database is available for comparative purposes.

NMR is a useful technique to study protein dynamics. Unlike X-ray diffraction, the method used to study crystallized protein structure, NMR can capture the state of a protein at many different solvent conditions and temperatures. In this way, the route of unfolding and refolding, kinetics and dynamics of any protein may be assessed. In the literature 90% of all NMR studies have been carried out by probing ^{1}H and ^{13}C nuclei, which are easy to detect and give immense details about protein dynamics. Herein, ^{1}H and ^{13}C signals of protein were examined under different environmental conditions (solvents, pH, temperature, etc.) in hopes to observe and to correlate changes of signal against structure. Water/deuterium mixtures were used to maintain signal lock. In many cases, water suppression was also employed to simplify protein spectra. Most experiments were arrayed in temperature and/or time to promote fast data collection and analysis. The effect of temperature, pH and several chemical substances were analyzed. In addition surface-exposed side-chains were ^{13}C-labeled, permitting a greatly simplified analysis of protein surface dynamics.

To better perform the simulations, a small High Performance Computing cluster was assembled and, namely, NAMD, a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems, was used [Kalé *et al.*, 1999]. NAMD was chosen because of its feature to readily implement all necessary MD methods such as CHARMM forcefields and temperature coupling methods. Moreover, it has no limit on the size of the protein to be simulated and is free to academic users. Heat capacity, fluctuations and relaxation of Cα atoms, radius of gyration can be calculated and compared with experimental findings.

In the following section, a brief outline of MD theory is given, the quantities computed from the MD results are described, and the experimental techniques used are outlined. In chapter 3, MD methods are introduced and simulation results are analyzed. Change of polar surface area is given and the relationship between heat capacity and radius of gyration is discussed. Theoretical solution scattering spectra are given for each temperature, providing a basis to speculate on the extent of unfolding *in silico*. In chapter 4, ^{1}H NMR results at different temperatures and pH values, natural abundance ^{13}C experiments and relaxation experiments with isotopically labeled proteins are presented. The protective effect of thiamine on protein stability and gelation is analyzed. Concluding remarks are made in Chapter 5, and possible future directions are outlined.

2. THEORY

2.1 Lysozyme

"We shall hear more about lysozyme"
Sir Alexander Fleming

Sir Alexander Fleming, also known for his discovery of penicillin, was suffering from a cold. A drop of mucous from his nose fell onto an agar plate, causing colonies on the plate to fade. He concluded that a lytic substance in the mucous was responsible for this effect, and lysozyme was discovered. Lysozyme is a small enzyme found in egg white, tears, and other secretions. It attacks the protective cell walls of bacteria. Bacteria normally possess a tough skin of carbohydrate chains, interlocked by short peptide strands, which brace the membrane against high osmotic cellular pressures. Lysozyme breaks the peptidoglycan carbohydrate chains in the walls by hydrolyzing the bond that connects N-acetylmuramic acid with the carbon atom at position four of N-acetylglucosamine, destroying the structural integrity of the cell wall. Subsequently the bacteria burst under their own internal pressure [Fleming, 1922].

Since Fleming's discovery of lysozyme, this enzyme has constantly been a subject of discussion and application in the life sciences. When the term lysozyme is loosely used, hen egg-white lysozyme is generally meant: It is the classical representative of this enzyme family. Hen egg-white lysozyme is remarkable in many ways. It was the first protein, which was sequenced and found to contain all the twenty usual amino acids [Canfield, 1963; Jolles and Jolles, 1961]. It was the second protein and the first enzyme structure to be solved by X-ray diffraction methods (the first was Myoglobin, solved in 1960 by John Kendrew at Cambridge) by the team led by David Phillips [Johnson, 1998; Blake *et al.*, 1965].

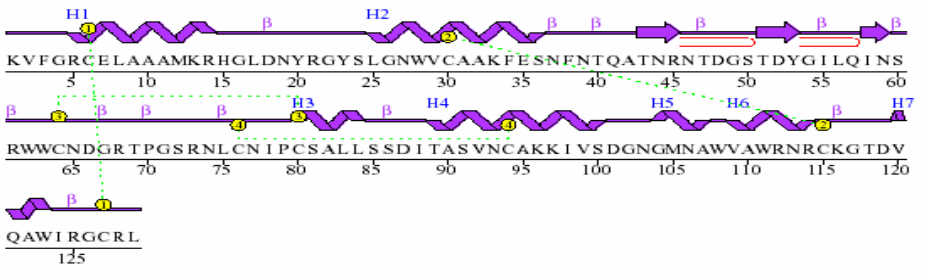


Figure 2.1: Promotif plot of HEW lysozyme secondary structure: 1, beta sheet; 2, beta hairpins; 3, strands; and 4, disulphide.

2.2 Molecular Dynamics

"Time is defined so that motion looks simple"
John Archibald Wheeler

Molecular dynamics, the science of simulating the motions of a system of particles [Karplus and Petsko, 1990], is a computer technique where the time evolution of a set of interacting atoms is monitored and followed by integrating their equations of motion. As a counterpart to experiment, MD simulations are used to estimate equilibrium and dynamic properties of complex systems that cannot be calculated analytically. Representing the exciting interface between theory and experiment, MD occupies a significant position at the crossroads of mathematics, biology, chemistry, physics and computer science [Schlick, 2002].

Molecules continuously interact among themselves and with their environment. Their dynamic motions can assume a wide range of thermally accessible states of a system, thereby providing as basis to correlate molecular structure and function. By following the dynamics of a molecular system in space and time, information concerning structural and dynamic properties such as geometry, energy, fluctuations and perhaps large scale deformations of proteins can be obtained [Schlick, 2002].

French mathematician Pierre Simon de Laplace (1749-1827) recognized the far reaching implications of Newtonian physics. He envisioned the possibility to predict future states by animating Nature's forces [Schlick, 2002], well known as Laplace's vision. In time, Schrödinger proposed a molecular mechanism for the organization of living phenomena, i.e., the 'order-from-order' mechanism, where ordered dynamics are composed

of packets of molecular dynamics having order, just as the ordered dynamics of a watch is caused by orderly movements of its smaller mechanical elements [Schimizu, 1979].

The first MD simulation of a protein was done in 1977 [McCammon *et al.*, 1977], but it took some time before more complicated simulations were possible. While at the time simulations were limited between a few ps to 25ps, Lewitt managed a 132ps simulation of bovine pancreatic trypsin inhibitor (BPTI). He found that hydrogen bonds are variable, that many break and reform again during the simulation, that the overall solvent-accessible area remains close to that of the X-ray structure, and that polar charged residues become less solvent exposed while non-polar hydrophobic residues become more solvent exposed. Together, these results provided a conceptual model of protein dynamics in which the molecule was envisaged to typically vibrate about a particular conformation, but also to suddenly change conformation periodically in the course of jumping over an energy barrier into a new region of conformational space [Lewitt, 1983].

Nowadays, all-atom MD simulations of protein-in-solvent systems are able to realistically map the protein-unfolding pathway. Their strong correlation with folding-unfolding experiments suggests that simulated unfolding events may also shed light on folding [Mayor *et al.*, 2000].

In addition to the all-atom MD method, one can also simulate partial dynamics of proteins to interpret the motions of proteins and the conformational transitions that potentially play a role in protein folding. Daggett, Kollman and Kuntz performed a long molecular dynamics simulation of polyalanine at high temperature [Daggett *et al.*, 1991]. Using this approach, they obtained a description of the overall structure and inherent flexibility of the chain as well as a structural picture of the conformational changes that occur. In this way, both equilibrium properties of the peptide, and dynamics & mechanisms of structural transitions can be addressed.

The solvent effect is also considered in MD. The dynamics of water in foods and several model systems have been studied. Correlation times of water in such model systems and the combination of molecular dynamics computations with NMR relaxation techniques have been discussed by the Baianu group located at the University of Illinois [Baianu *et al.*, 1991].

In the history of MD simulations, the year 1995 was marked by the immediate

effects of fast Ewald methods, an accomplishment resulting from several years' work in the search for an adequate treatment of the electrostatic long-range forces [Louise-May *et al.*, 1996].

To explore the dynamics of proteins using T4 lysozyme as a model, collective motions and interresidual correlations in proteins were studied using low-resolution simulations, with knowledge-based potentials [Bahar *et al.*, 1997]. Trajectories were partitioned into modes, and the slowest ones were analyzed to elucidate the dominant mechanism of collective motions. There appeared to be a correlation between groups involved in highly cooperative motions, as revealed by simulations, and highly protected regions during unfolding, as measured by pulsed H/D exchange and 2-D NMR experiments.

Fersht and Daggett modeled BPTI in solution using MD simulations at a variety of temperatures to describe unfolding characteristics of proteins. BPTI unfolded at high temperature, assuming an ensemble of conformations with all the properties of the molten globule state [Fersht and Daggett, 2002]. They outlined the structural details of the actual unfolding process between the native and molten globule states. The first steps of unfolding involved expansion of the protein, which disrupted packing interactions. The solvent-accessible surface area also quickly increased. The unfolding was localized mostly to the turn and loop regions of the molecule, leaving the secondary structure intact. Then, there was more gradual unfolding of the secondary structure and non-native turns became prevalent. This same trajectory was continued and more drastic unfolding occurred, resulting in a relatively compact state devoid of stable secondary structure.

2.2.1 MD Theory

In MD, Newton's 2nd law (below) and classical mechanics are applied:

$$F_i = m_i a_i \tag{2.1}$$

to each atom i in a system of N atoms. Here, m_i is the atom mass, $a_i=d^2r_i / dt^2$ is its acceleration, and F_i is the force acting upon it due to the interactions with other atoms. The force acting on particle i can also be expressed as the gradient of the potential energy, V:

$$F_i = -\Delta_r V \tag{2.2}$$

Combining the equations (2.1) and (2.2) yields:

$$\frac{-dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \tag{2.3}$$

where the potential governing the motion of particle i is the sum over all the effective interactions, and is called a forcefield [Leach, A. R., 1996]. Newton's equations of motion can then relate the derivative of the potential energy to the changes in position as a function of time. The force field used in NAMD is the CHARMM [Brooks *et al.*, 1983] forcefield which includes 2-, 3-, and 4-body interactions, electrostatic interactions, and van der Waals interactions [Kalé *et al.*, 1999].

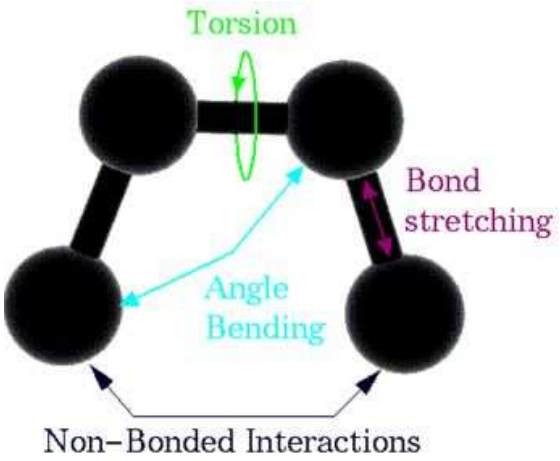


Figure 2.2: Representation of intermolecular potential forces.

The computations involved in each time step of MD can be broken into several portions. First, the forces acting on each atom are computed according to the empirical forcefield that approximates intermolecular forces. Since potential is a function of the atomic positions of all the atoms in the system, and due to the complicated nature of this function, there is no analytical solution to the equations of motion. Rather, once these forces are computed, a numerical integration scheme is used to update the positions and velocities of all the atoms. Solving a dynamic system means determining how to project the

system forward in time from some set of initial conditions, i.e., computing future positions as a function of time. For instance given the acceleration, an approximate velocity for the atom can be computed for a given period of time and changes in atomic coordinates can be determined.

Several methods are available for numerical integration. In NAMD, the Verlet method is implemented. At a given time step, the force on an object is specified and the position is desired as a function of time. Since only accelerations are specified, one integration must be done to calculate the velocity and a second done to calculate the position. Accordingly, the Verlet scheme first updates the position, and then uses the old and new positional information to update the velocity [Haile, 1992].

$$X_n = X_{n-1} + V_{n-1} \cdot \Delta t + \frac{1}{2} \cdot \alpha_{n-1} \cdot \Delta t^2 \tag{2.4}$$

$$v_n = v_{n-1} + \frac{1}{2} \cdot (\alpha_n + \alpha_{n-1}) \cdot \Delta t \tag{2.5}$$

The choice of the integration time step and the simulation time are very important. The time step chosen should allow the fastest motion in a molecule to be observed [Brooks *et al.*, 1983]. Since MD applied to biological macromolecules predicts the fluctuations in the relative positions of the atoms in a protein, knowledge of these motions provides insight into biological phenomena, such as the role of flexibility in ligand binding. Taken to another level, such insight can presumably help to unravel pathways of unfolding [Fersht and Daggett, 2002; Fersht and Daggett, 2003].

Event	Time
Bond stretch	~1 to 20fs
Elastic domain modes	100fs to several ps
Water reorientation	4ps
Inter-domain bending	10ps to 100ns
Globular protein tumbling	1 to 10ns
Aromatic ring flipping	100μs to several sec
Allosteric shifts	2μs to several sec
Local denaturation	1ms to several sec

Table 2.1: The characteristic times for common molecular events [Brooks *et al.*, 1999]

A MD simulation generates a sequence of points in phase space as a function of time, where these points belong to some ensemble. These correspond to different conformations of the system and respective momenta. In simulations, the Canonical Ensemble (NVT) is used, i.e., a collection of all systems whose thermodynamic state is characterized by a fixed number of atoms, *N*, a fixed volume, *V*, and a fixed temperature, *T*.

The equilibration of systems to make the relevant properties converge to their equilibrium values is an important factor. Simulation time depends on the size of the system and computer power. While short simulation times may not reflect the protein behavior at that temperature, quite long simulation times are beyond currently available computational power. The periodic boundary conditions method was developed to make a simulation behave as if it was infinite in size, effectively removing the effects of the surface of a finite sized system. This strategy also ensures that the internal structure of the sample is dominated by surface rather than bulk forces [Allen and Tildeslay].

Before any equilibration, the energy of the system must be minimized to remove any bad contacts in structure, which cause some atoms to have very high potential energies. The Steepest-Descend (SD) method is used to minimize protein-solvent system.

2.2.2 Heat Capacity

Heat capacity under constant volume, C_v, is defined as that heat quantity which is required to increase the temperature of a unit mass of a material by 1K. Heat capacity is a fundamental property of any material and can be calculated by differential scanning calorimetry (DSC) as it is a temperature dependent quantity. Heat capacity may be theoretically obtained from the energy output of MD simulations by:

$$\frac{\left\langle \left(E - < E >\right)^2 \right\rangle}{k_b T^2} \tag{2.6}$$

where E is the potential energy at a given time, $<E>$ is the average energy value, k_B is the Boltzmann constant, and T is temperature.

2.2.3 Fluctuation Vector

There have been a number of advances in atomic resolution simulations of biomolecules [Fersht and Dagget, 2002]. These have arisen partly from improvements made to computer power and partly from algorithmic improvements. There have also been advances in measuring time-dependent fluctuations in proteins. In one study, the time-scale for carbonyl fluctuations about the Cα-Cα axis and kinetic rates for cation movement in the channel was found to be equivalent, suggesting a correlation between molecular dynamics and kinetics [Cross *et al.*, 1999].

In the experiments, a fluctuation vector attached to the Cα atoms of the trajectory is first computed by splitting trajectories into 400ps time pieces. Then, best-fit superpositions are taken from these pieces. Average structure is computed and the trajectories are once more best-fitted to this average structure. In this manner the resulting trajectory reflects contributions from the motions of the internal coordinates only [Baysal and Atilgan, 2002].

2.2.4 Relaxation Phenomena

This phenomena characterizes the motion of the fluctuation vector by a relaxation function of time and temperature, $C(t,T)$, namely:

$$C(t,T) = \frac{\overline{< \Delta R(0,T) \cdot \Delta R(t,T) >}}{\overline{< \Delta R^2(T) >}} \tag{2.7}$$

Usually $C(t,T)$ will have contributions from many different homogeneous processes with different relaxation times, and their collective effect on relaxation will be observed as heterogeneous dynamics [Deshenes and Vanden Bout, 2001].

2.3 Nuclear Magnetic Resonance

"Every great advance in science has issued from a new audacity of the imagination."
John Dewey

2.3.1 Overview

Nuclear Magnetic Resonance (NMR) has become a general tool of structural and dynamic analysis during the last few decades. NMR studies of protein dynamics, in principle, give insight into the relation between motion and function [Smith].

The first experiments of C.J. Gorter were conducted in 1936. Although his attempts to observe the resonance property of nuclear spin in the presence of a magnetic field were unsuccessful, he brought attention to the potential of resonance methods [Gorter, 1936]. In 1945, Harvard University professor Edward Mill Purcell and students Torey and Pound assembled a radiofrequency apparatus that detected the energy transition between nuclear spin states. Using about 1 kg of paraffin wax, an absorbance was predicted and observed. At the same time, Felix Bloch at Stanford University also demonstrated the phenomenon of NMR. Their discovery opened the way to the development of analytical spectrometry methods for the determination of the chemical structures of unknown organic compounds. Bloch and Purcell were awarded the Nobel Prize for Physics in 1952 for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith.

In the 1960s, several groups began to apply the tools of NMR spectroscopy to biological molecules. The first protein spectra were difficult to interpret, particularly due to poor resolution, and it was not obvious from the outset that NMR could actually be used to determine protein tertiary structure. It was suggested that interactions between amino acids in folded proteins yielded in the observed broad peaks in these spectra, but this point remained in question for many years.

Afterwards, they made definite improvements. Observation of chemical shifts [Purcell *et al.*, 1946] showed the potential usefulness of NMR as an analytical method [Bloch *et al.*, 1946]. Albert Overhauser's theory, namely, the Nuclear Overhauser effect, was introduced and became established as a very important tool for the determination of complex molecular structures. The Fourier Transform pulse NMR technique was

introduced by Ernst, and 2D NMR techniques were introduced by Jeener. Clearly, NMR expanded well beyond its use in analytical chemistry and became an important tool for structural analysis of biological macromolecules. The basis of NMR is also used in medicine. Magnetic resonance imaging (MRI), for instance, is founded on the same principles.

The first experimental study of a refolding pathway in the literature dates back to 1963. A Russian team focused on understanding reversible unfolding of ribonucleoprotein strands using packing models. In 1975, the first reversible unfolding study of lysozyme, based on spectral measurements, was published [Elwell and Schellman, 1975]. In this study, spectral properties of T4 lysozyme were determined and these properties were used to follow the unfolding transition. The reported ultraviolet absorption spectrum and solvent perturbation difference spectrum indicated that the aromatic amino acids were extensively exposed to solvent. The group also determined reversible thermal denaturation conditions of lysozyme at acidic pH.

It was later found that the folding process could not be explained by a simple two-state mechanism. Rather it involved intermediate forms that had the same elements of secondary structures as the native form and the transition between the intermediates and the fully denatured states was extremely rapid. Consequently, new techniques had to be applied to study rapidly transcending, unobservable intermediate forms [Nozaka *et al.*, 1978].

A year later, ^{13}C NMR spectra of ribonuclease A revealed many details about reversible unfolding over the pH range 1-7 and between 6˚C-70˚C. Evidence was presented for the existence of intermediate unfolding stages, commencing at a minimum of 10˚C difference on either side of the main unfolding transition, particularly at low pH values. These claims were supported by measurements of spin-lattice relaxation times (T_1) and of Nuclear Overhauser enhancement. It was also shown that the native protein has more variability of structure at low pH than at neutral pH, and also interchanges more rapidly with the semi-structured, denatured state [Howarth, 1979].

The first examples of reversible conformational changes in proteins have been monitored using ^{1}H-NMR, with glycophorin as the model. It was found that the helix of the hydrophobic domain is remarkably resistant to conventional denaturing conditions including pH and temperature extremes [Cramer *et al.*, 1980].

In another NMR study, the reversibility of ribosomal protein E-L30 was shown to involve a fast and a slow equilibrium, which depended on the degree of protonation of histidine residues [van de Ven, 1987]. Both the fast equilibrium between protonated and deprotonated histidines and the slow equilibrium between folded and unfolded protein were monitored by means of 500-MHz ^{1}H NMR spectroscopy. A greater protonation of histidine residues appeared to be detrimental in the unfolding process of the protein. It was shown, however, that even when the histidines are uncharged, the protein has only limited stability.

2.3.2 Magnetic Properties of Nuclei

"Anyone who says that they can contemplate quantum mechanics without becoming dizzy has not understood the concept in the least."
Niels Bohr

Electrons, neutrons and protons, the three particles which constitute an atom, have an intrinsic property called spin. This spin is defined by the fourth quantum number of any given wave function, obtained by solving the relativistic form of the Schrödinger equation (SE). It represents a general property of particles, which is often described using electrons as the model system. In atoms, electrons circulate around the nucleus, generating a magnetic field. The generated field has an angular momentum associated with it. It turns out that there is also an angular momentum with the electron particle itself, denoted as the spin, and this gives rise to the spin quantum number, m_s.

Spin angular momentum is quantized and can take different integer or half-integer values depending on what system is under study. If the relativistic SE for the electron is solved, the values +½ and -½ are obtained. Since the Pauli principle states that no two species can have the same four quantum numbers, only antiparallel electrons can reside in the same atomic orbital [Arnold et al, 1951].

Like the electron, protons and neutrons also have a spin angular momentum which can take values of +½ and –½. In the nucleus, protons can pair with other antiparallel protons, much in the same way as electrons pair in forming a chemical bond. Neutrons do the same. Paired particles, with one positive and one negative spin state have a net spin of zero. A nucleus with unpaired protons and neutrons will have an overall spin, with the

number of unpaired elements contributing ½ unity to the overall nuclear spin quantum number, I. When the sum is non-zero, a nucleus will have a spin angular momentum and an associated magnetic moment, defined by the direction and magnitude. It is this magnetic moment that is exploited in modern NMR experiments.

2.3.3 The Chemical Shift

The electron density around each nucleus in a molecule varies according to the types of nuclei and bonds in the molecule. The opposing field and therefore the effective field at each nucleus will vary. This is called the chemical shift phenomenon.

The chemical shift value of a nucleus is the difference between the resonance frequency of the nucleus and a standard, relative to the standard. This quantity is reported in parts per million (ppm) and given the symbol delta, δ.

$$\delta = \frac{(\nu - \nu_{REF}) \times 10^6}{\nu_{REF}} \tag{2.9}$$

The chemical shift is a very precise metric of the chemical environment around a nucleus. For example, the hydrogen chemical shift of a CH_2 hydrogen next to a Cl atom will be different than that of a CH_3 next to the same Cl atom. It is therefore difficult to give a detailed list of chemical shifts in a limited space, but there are predictive methods in use.

The existence of chemical shift dispersion is crucial for the application of NMR spectroscopy to biomolecules, but the direct interpretation of shift tensors in terms of structure and dynamics is often difficult. Proton shifts reflect environmental influences from nearby aromatic groups, metal sites or hydrogen-bonding partners. Shifts for carbon and nitrogen generally reflect local bonding interactions, often in ways that would allow the local structure to be inferred. The anisotropy of the shielding tensor is also of interest. It influences the resonance position in partially-ordered samples and has major consequences on spin relaxation processes, even in isotropic systems [Case, 1998].

2.3.4 Spin-Lattice & Spin-Spin Relaxation

NMR spin-relaxation can provide unique insight into overall & internal motions, as well as the time-dependence of conformational fluctuations, especially on the picosecond to nanosecond time scale. As mentioned previously, these events can also be probed by simulations. A great deal has been learned from such simulations, about the general nature of such motions, and their impact on NMR observables [Case, 2002]. In principle, relaxation measurements should also provide valuable benchmarks for judging the quantitative accuracy of simulations. There are two types of relaxation to consider.

The first process, called population relaxation, refers to nuclei that return to the thermodynamic state in the magnet. This process is also called T_1 relaxation, where T_1 refers to the mean time for the bulk magnetization vector to return to its equilibrium state. Once the population is relaxed, it can be probed again, since it is in the initial state.

The precessing nuclei can also fall out of phasal alignment with each other (returning the net magnetization vector to a nonprecessing field) and stop producing a signal. This is called T_2 relaxation. In this state, the population difference required to give a net magnetization vector is not at its thermodynamic state. Some of the spins were flipped by the pulse and will remain so until they have undergone population relaxation. T_1 is always larger (slower) than T_2.

2.3.5 Nuclear Overhauser Effect

The Nuclear Overhauser effect (NOE) is a phenomenon whereby the polarization of one spin population may be used to enhance the polarization of a second spin population to which it is coupled by dipolar interaction. Originally conceived as a method for increasing the sensitivity of NMR by the transfer of polarization from coupled electron spins, NOE is more usually observed between coupled nuclear spins. Experimentally, the more sensitive nucleus (usually the proton) is irradiated by a B_1 field, which causes saturation. This enhances the polarization of the coupled spin, thus increasing the signal from this nucleus.

In ^{13}C NMR, the NOE effect is utilized to significantly enhance the normally weak

^{13}C NMR signal of groups such as methylene carbons. The sample is irradiated with broad-band radiation to saturate all attached proton. Spin-lattice energy transfer occurs to the neighboring carbons and the NMR signals are greatly enhanced. The use of this technique allows ^{13}C NMR spectra to be collected in a relatively short time, but, since the intensity of the NOE depends on the number of attached protons and other possible mechanisms for spin-lattice energy transfer, the resulting integration is no longer proportional to the number of ^{13}C nuclei giving rise to the signal; i.e., in such experiments the integration of spectra is essentially meaningless unless the experiment is carefully calibrated.

2.4 Solution Scattering

Solution scattering is an effective technique to determine the low-resolution, basic shape of molecular complexes, providing structural information that reveals global conformation changes in biomolecules and protein folding in solution.

In a scattering experiment, a solution of macromolecules is exposed to X-rays or thermal neutrons. The scattered intensity, $I_{(s)}$, is recorded as a function of the momentum transfer s;

$$s = \frac{4\pi \sin\theta}{\lambda} \tag{2.8}$$

where 2θ is the angle between the incident and scattered radiation and λ is the wavelength. Typically, the solvent scattering is subtracted. The random positions and orientations of particles result in an isotropic intensity distribution, which, for monodisperse solutions of noninteracting particles, is proportional to the scattering from a single particle averaged over all orientations. The net particle scattering is proportional to the squared difference in scattering length density between particle and solvent [Svergun and Koch, 2002]. This approach also holds great potential for resolving the structural kinetics of macromolecular complexes using time resolved solution scattering to produce low-resolution movies of events, such as the assembly and operation of molecular machines [Svergun, 1995].

3. MD STUDY OF HEW LYSOZYME

"All science is either physics or stamp collecting"
Ernest Rutherford

3.1 Overview

To simulate the unfolding of HEW Lysozyme, different MD simulations were prepared and run at 5K intervals between 300K and 355K. The energy and temperature output, as well as the coordinates for each time step were analyzed by algorithms coded in-house and ready tools cited below were used for visualizing methods. These trials were repeated to enhance accuracy. Because lysozyme unfolding is much longer than the ns time scales examinable by MD, a 4ns run at 500K was done, following the methodology of Daggett *et al.* whereby unfolding has been shown to occur at much shorter time scales at very high temperatures [Fersht and Daggett, 2002]. Moreover, MD runs at 20K intervals between 150 – 300K have been carried out to compare with earlier results on the glassy behavior of proteins.

3.2 Computational Systems and Computer programs

For the MD simulations, in-house computer Clusters were used. Two PCs with two Xeon 2500Mhz CPUs connected by MPI and SSH protocols were utilized. An experimental small cluster of 10 PCs with Pentium III 1000 MHz CPUs, (MDBF-CLUSTER) was designed and installed. A Linux box with two Xeon 3000MHz CPUs (Render) was used for single long runs.

Dynamics simulations were realized using the package Not Another Molecular Dynamics (NAMD), a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems. Charm++, developed by Proffesor Kale and co-

workers [Kalé *et al.*, 1999], simplified parallel programming and provided automatic load balancing, features which are crucial to the performance of NAMD. Based on Charm++ parallel objects, NAMD scales to hundreds of processors on high-end parallel platforms, and tens of processors on commodity clusters using gigabit Ethernet. NAMD is file-compatible with AMBER, CHARMM, and X-PLOR and is distributed free of charge with source code. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign [Kalé *et al.*, 1999].

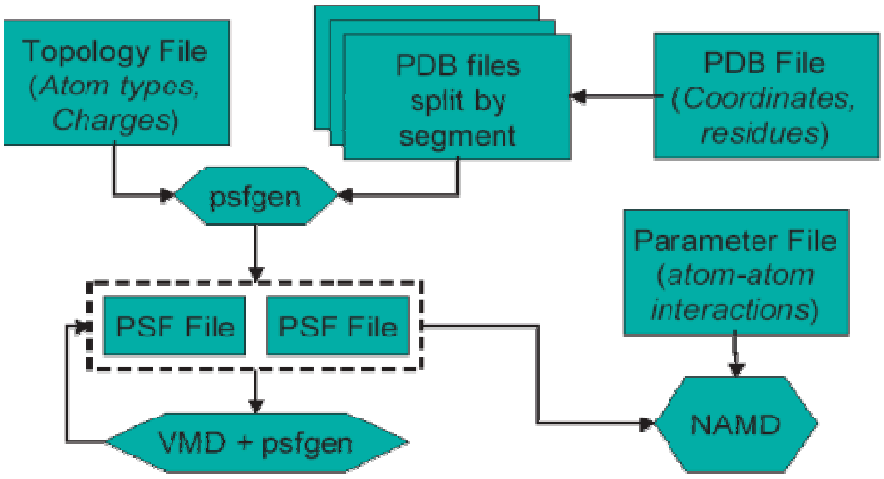


Figure 3.1: Molecular dynamics flowchart with NAMD code.

Molecular Visualizations and some of the analysis were done by Visual Molecular Dynamics (VMD), a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in tcl and python scripting. VMD supports computers running MacOS-X, UNIX, or Windows, is distributed free of charge, and includes source code [Humphrey *et al.*, 1996].

Evaluation of X-ray Solution Scattering curves are calculated by Crysol using the average coordinates of in-house simulation results. The program uses multipole expansion of the scattering amplitudes to calculate the spherically averaged scattering pattern and takes into account the hydration shell [Svergun *et al.*, 1995].

3.3 System Preparation

The Hen Egg White Lysozyme crystal structure file 6lyz.pdb [Diamond, 1974] was taken from the PDB.

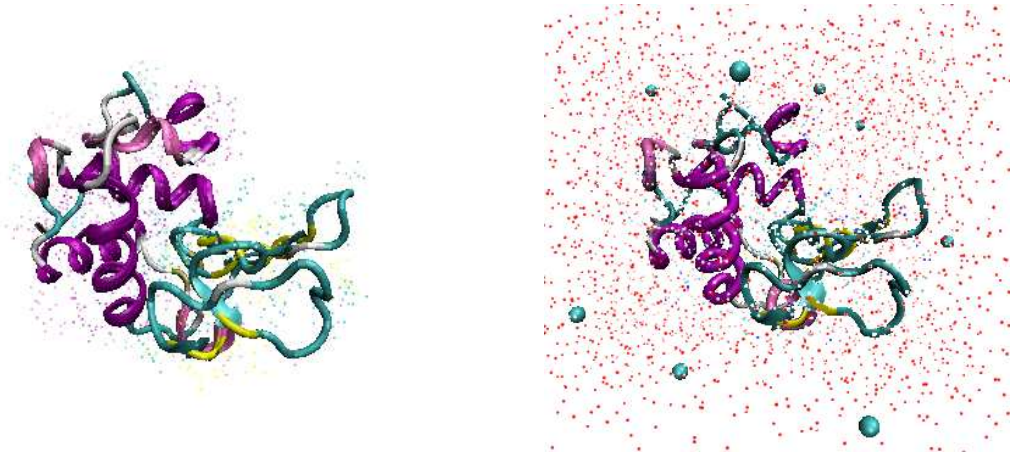


Figure 3.2: Hen egg-white lysozyme (6lyz). Left, crystal structure display; right, after adding a water box and chloride ions for neutralization. Chloride ions are enlarged for a improved viewing.

Hydrogen atoms do not appear in this structure file, since their sizes are too small to interact effectively with X-ray radiation.

For the purpose of obtaining a better simulation, in vivo environments are being mimicked as closely as possible. In accordance with this theme, a new lysozyme structure with the coordinates of missing hydrogen atoms were created. This structure is solvated in a box containing TIP3 waters. The TIP3 model is suggested to be a better choice for the development of a balanced force field. It is expected that water models, which include electronic polarization, will allow for better modeling of pure solvent properties [MacKerell, 2001].

With the solvate module in VMD, a box for PBC was created where the thinnest layer of water was fixed at 5Å. Each side of the box was 54Å.

When using periodic boundary conditions, the energy of the electrostatic interactions were calculated with particle-mesh Ewald (PME) summation, which requires the system be electrically neutral. The autoionize plug-in of VMD, scripted by Ilya Balabin,

was used to make the net charge of the system zero by adding eight chloride ions to the solvent. The plug-in distributed the ions randomly in the solvent, with some distance limitations maintained to create a better charge distribution. The final structure files containing eight chloride ions and lysozyme in the water box were used in the NAMD simulations.

3.4 Simulation Details

The system was first energy minimized to 2.5 X 10^{-4} kcal/mol/Å of the derivative by 5000 conjugate gradient iterations. The minimization took place in approximately 300s on this cluster. These new coordinates of the system were treated as the initial coordinates in all the MD simulations. All bonds of the protein and water molecules were constrained by the RATTLE algorithm [Andersen, 1983]. Integration was carried out by the velocity Verlet algorithm. The systems were equilibrated for 500ps with a time step of 2fs, while maintaining the temperature by direct velocity scaling. The data-collection stages were of 2ns length. At this stage a temperature coupling method used to control the temperature. Data was recorded every 2ps and each 200ps portion of the trajectories were treated as a separate sample. Thus 10 data sets at each temperature were collected between 150K and 355K, and the calculated quantities were averaged over this range.

3.5 Results

MD simulations were obtained for a 2ns time frame and temperatures between 150-375K at 5K increments. Each simulation followed a 500ps equilibration step. Unfolding experiment to observe the folding process of lysozyme is done at 500K for 4 ns following a 1ns equilibration.

3.5.1 Heat Capacity

Constant volume heat capacity was calculated from the energy output of simulations.

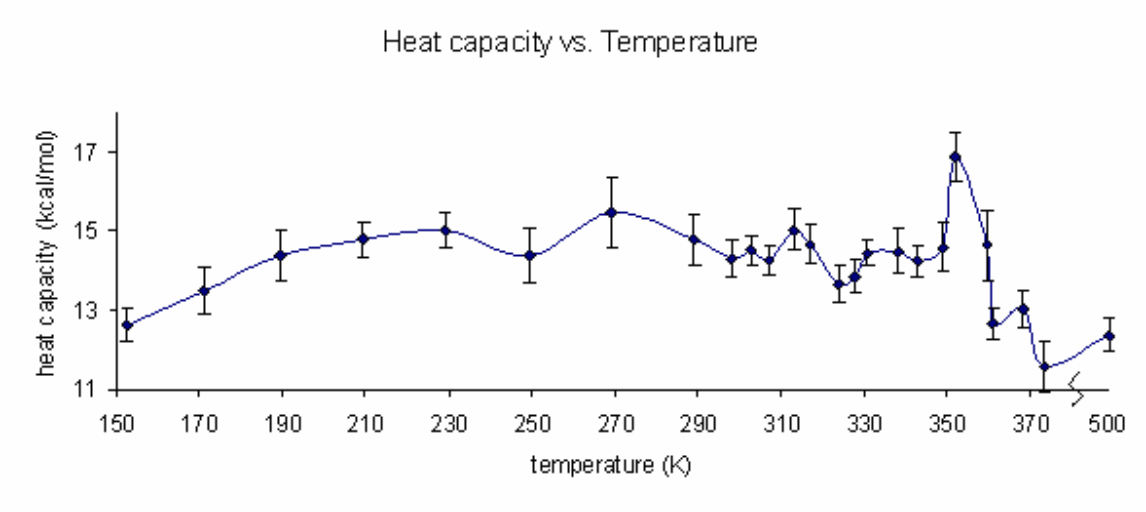


Figure 3.3: Heat capacity versus temperature. At each simulated temperature, 10 heat capacity values are calculated for each 200ps and mean values are plotted with standard errors. Last value belongs to 500K added close for clarity.

In these simulations, the initial increase of heat capacity, beginning at 150K and ending at approximately 210K, is associated with the so-called "protein glass transition". C_v remains constant thereafter until 330K. A sharp peak is observed around 350K, signifying the onset of unfolding.

The interpretation of the change in C_v as a function of temperature is complicated. There are two main contributors related to temperature-induced unfolding in particular:

i. Structural changes associated with the unfolding of the chain influencing the hydration; Changes in hydration that are associated to chain unfolding; and
ii. Positional fluctuations of the atoms.

A model depicting both of these effects is presented in figure 3.4 [Cooper, 2000].

Simulations, on the other hand, cannot probe the effects of heat capacity changes the time intervals are too short to observe larger structural changes associated with unfolding of chain segments. Information related to temperature-induced fluctuations, however, is contained in the calculations, as shown by the peak spanning the 355-365K region (figure

3.3). A DSC trace depicting protein melting (figure 3.4) supports this finding in the sense that heat capacity of lysozyme changes most dramatically in the same temperature range.

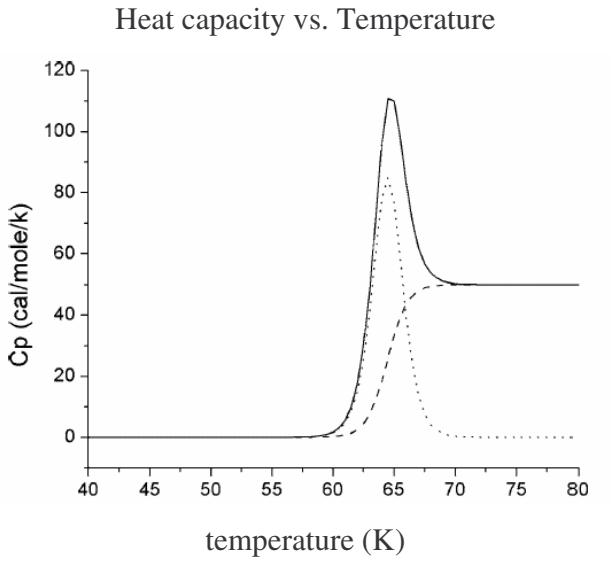


Figure 3.4: DSC trace of protein melting [Prabhu, N. H., and Sharp, K.A., 2005]. The dotted line shows temperature-induced fluctuations of heat capacity, the dashed line shows heat capacity changes in proceeding from native to denatured state, and solid line is their sum

3.5.2 Radius of Gyration

One indicator of unfolding is the onset of a change in the radius of gyration (R_g) with increased temperature. Hirai and coworkers have recently reported the R_g of HEW lysozyme in the temperature range 13-84 ˚C using wide angle X-ray scattering at different pH values [Hirai *et al.*, 2004]. In particular, the reported R_g is 15 ±0.5 Å until the onset of unfolding. Depending on the pH value an approximate 3-4Å increase in R_g is reported, the change being smaller at lower pH values. The R_g values obtained in the calculations show a similar qualitative behavior.

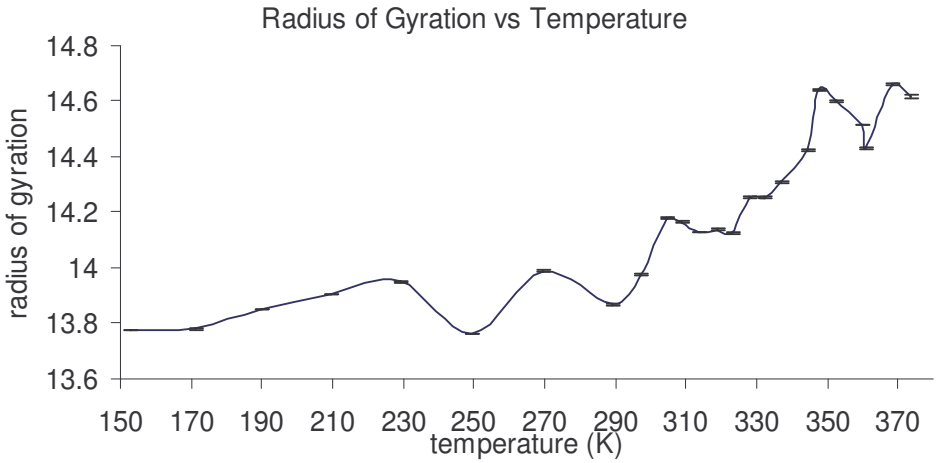


Figure 3.5: Radius of gyration is calculated at each simulated temperature and mean values are plotted with standard error.

The radius of gyration has a constant value of approximately 13.9Å in the temperature range 150-290K (figure 3.5). In the range of 300-330K, R_g increases to a constant value of 14.2Å. Another change in R_g was noted with an onset at 330K, with a dip in the value at the transition temperature, in agreement with C_v calculations, i.e., 355K. The increase of approximately 1Å in R_g, however, is much lower than what would be expected of large structural changes accompanying actual unfolding. Thus, the conjecture of the previous section is supported, namely, that the changes in C_v are mainly due to large scale fluctuations and not due to unfolding of structure. Since the simulation length of 2ns is too short to observe the actual unfolding process at these temperatures, the time evolution of R_g was also monitored during the 4ns, 500K simulations.

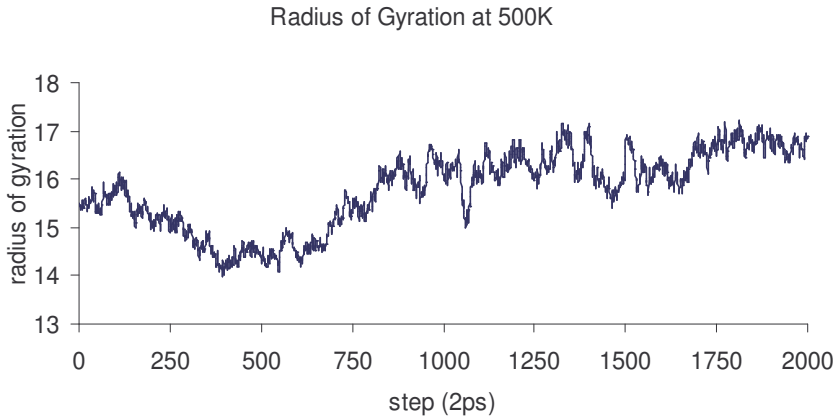


Figure 3.6: Radius of gyration curve at 500K.

The final structures have an R_g of 17Å, implying an increase of approximately 3Å. Thus, these structures may be good representatives of the unfolded structures, albeit the unrealistic simulation temperature, as shown by Fersht and Daggett [Fersht and Daggett, 2002]. They may also contain information on folding intermediates.

It has been shown that intermediate unfolding steps are observable even at extreme temperatures [Fersht and Daggett, 2002]. Water accessible surface area was calculated and compared with previous results. At 500K, water accessible surface area was in good agreement with the radius of gyration. This finding implied that each protein, with a characteristic distribution of amino acids and unfolding profile, can reveal a fingerprint in the early stages of unfolding.

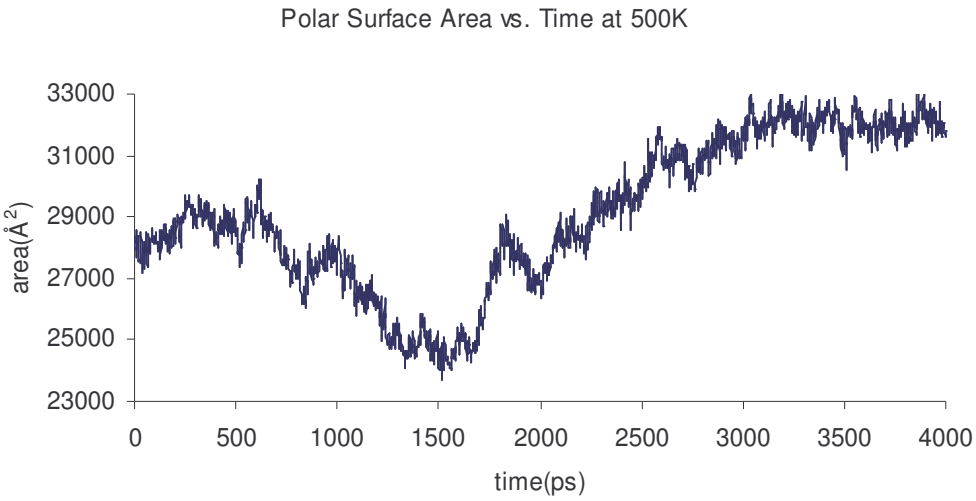


Figure 3.7: Polar surface area of hen egg-white lysozyme at 500K. Area decreases in the first 1.5ns and increases in the next 1.5ns, remaining constant thereafter.

3.5.3 Solution Scattering Functions

Small angle X-ray scattering (SAXS) is a potentially powerful method to obtain structural information from biological molecules in solution. The use of this technique in the laboratory has been limited by the long exposures necessary to obtain patterns on photographic film. To assess the extent to which structures are unfolded, the experimental data for the folded structures were compared, and corresponding data computed using the CRYSOL software [Svergun *et al.*, 1995]. Results at 150, 310, 345 and 500K are shown in figure 3.7.

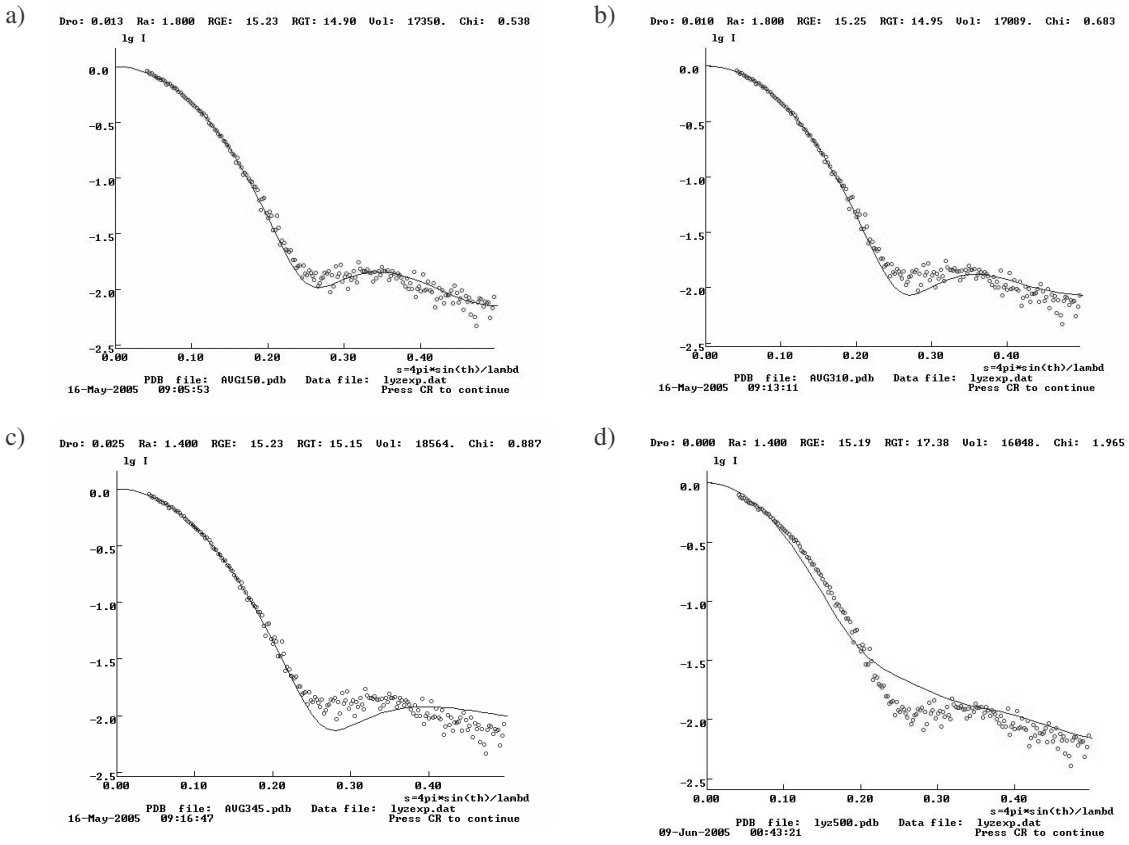


Figure 3.8: Solution scattering graphs of average structures of simulations for a) 150K b) 310K c) 345K and d) 500K.

It has been found that only the structures at 500K represent the unfolded state, supporting the findings of the previous section.

3.5.4 B-factors

Thermal fluctuations describing the reduction of the intensity of Bragg scattering due to motion of atoms about their equilibrium position (B-factors) are averaged over all residues. Such data treatment is useful to obtain information on the types of motions operating on the whole molecule or parts of it, and also to help derive general concepts of protein dynamics [Fraunfelder and McMahon]. B-factors of Cα atoms averaged over all residues are presented in figure 3.8. In the temperature range studied, three distinct regimes were observed. At low temperatures (up to 200K), the force constant characterizing the fluctuations of the protein is low, the force constant being directly proportional to the slope of the line passing through these data [Zaccai, G., 2000]. Upon the protein glass transition that occurs in the vicinity of 200K, the size of the fluctuations show an increased trend, characterized by a larger force constant.

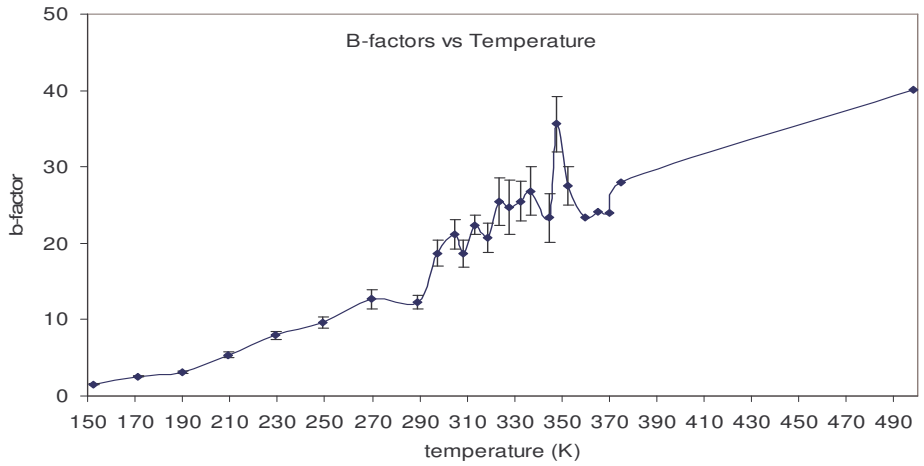


Figure 3.9: Thermal fluctuations for each simulation at different temperatures.

Finally, above approximately 300K, the force constant increases once more. These fluctuations eventually lead to the unfolding of the protein. From the previous studies, it was known that the details of the fluctuations (residue-by-residue) were the same at temperatures below approximately 320K for BPTI [Baysal, C., and A. R. Atilgan, 2005]. It has been found that the same trend also occurs for HEW lysozyme.

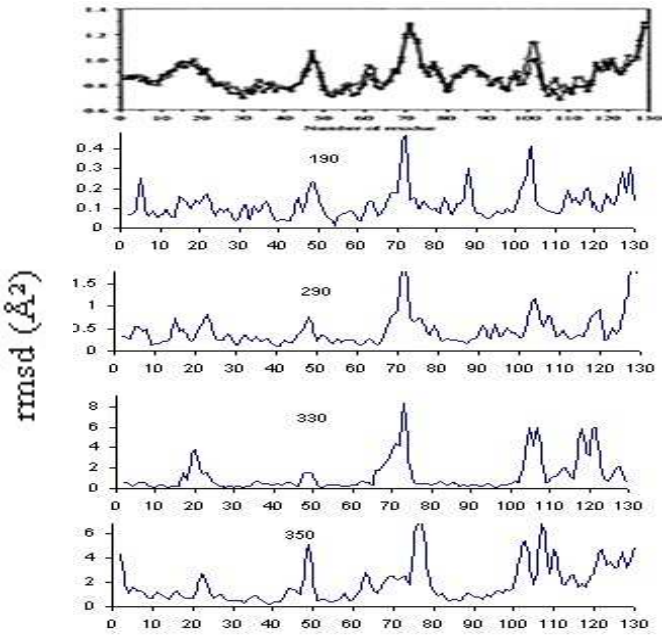


Figure 3.10: Experimental and computational fluctuations. Top figure is temperature factors of hen egg-white lysozyme. Below are residue-by-residue fluctuations at given temperatures.

Data of residue-by-residue fluctuations are shown in figure 3.8 at 190, 290, 330 and 350K, along with the experimental data from 1rfp.pdb [Motoshima, H., 1997]. At 500K, however

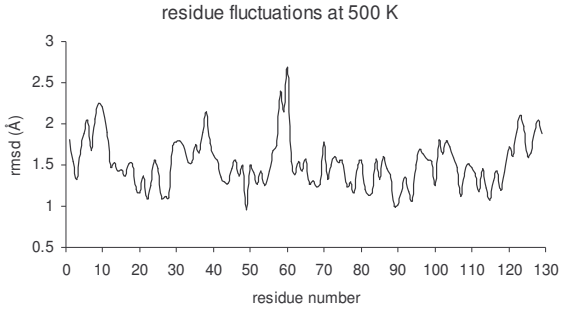


Figure 3.11: Residue-by-residue fluctuations of Cα atoms of HEW lysozyme at 500 K.

where unfolding was simulated, most fluctuated residues differ from the previous fluctuation results, because different events and forces such as domain bending and denaturation are expectedly taking place during unfolding.

3.5.5 Characterizing the Heterogeneous Dynamics

The motion of the fluctuation vector is characterized by a relaxation function of time and temperature for all the simulations. For simplicity, data is shown for only 5 different temperatures. The results nicely distinguish the protein dynamics at different temperature values.

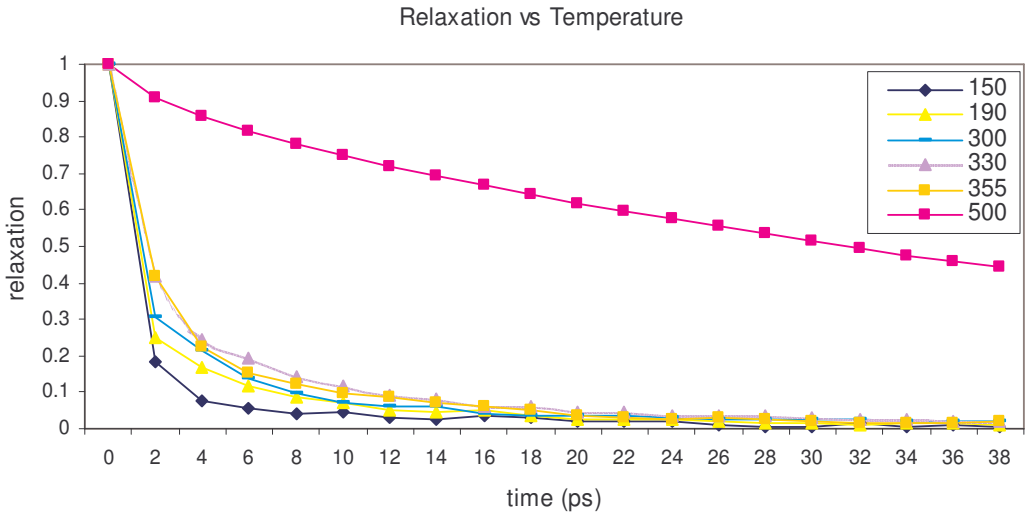


Figure 3.12: Relaxation function of hen egg-white lysozyme at several temperatures.

The stretched exponential fits of relaxation data are examined at all the temperatures by using equation 2.7. For a simple process characterized by a simple exponential decay, β happens to be 1. β tends to decrease from 1, not only if a larger number of contributing processes are at play, but also if these processes have time or length scales spanning different orders of magnitudes [Baysal and Atilgan, 2002]. The results show that the simulation data are well approximated by the fits at all time scales slower than 20ps. Similar to the results for BPTI, β increases from approximately 0.2 to 0.4 after the protein glassy transition, and remains at this value until approximately 330K, where it approximately increases further to a value around 0.5. At very high temperatures, e.g. 500K, β becomes much higher, coming close to 1, and representing the highly disordered unfolded state.

Note the remarkable resemblance between β and C_v gives a clue that the former may be a thermodynamical parameter.

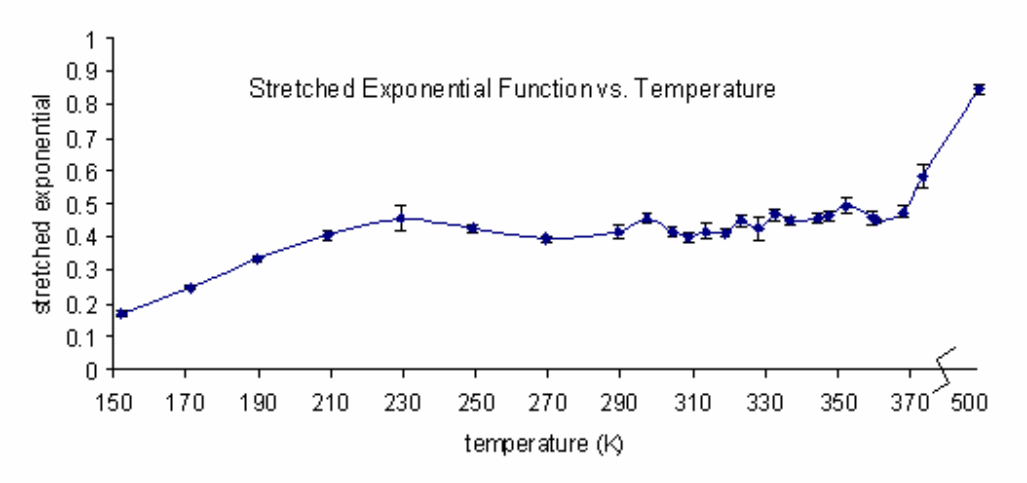


Figure 3.13: Stretched-exponential fit of relaxation data as a function of temperature

3.5.6 Unfolding Pathways

To characterize the unfolding pathways, the trajectories at 500K were monitored. Snapshots of the protein along these trajectories are shown in figure 3.14. Four amino acid residues (arginine 68, threonine 69, glycine 71 and serine 72) observed around the proline residue at position 70 were fluctuating more after 320K. At critical temperatures, where the onset of unfolding is assumed to occur, hydrogen bonding between Arginine 68 and Aspartic acid 48 was broken, causing a big conformational change in the protein.

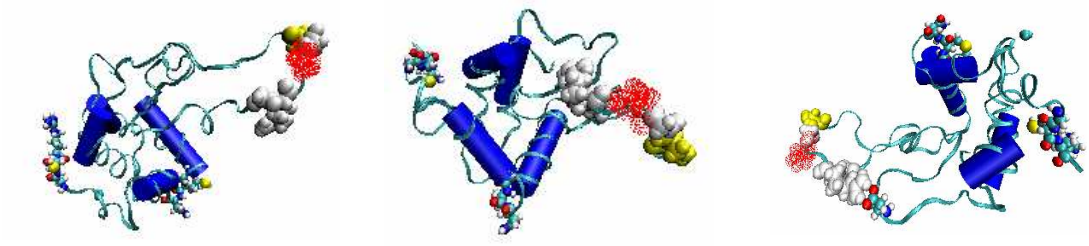


Figure 3.14: Graphical representations of most active residues. The red cloudy residue is proline at position 70. The left bottom region of the 3rd representation shows hydrogen bonding between residues 48 and 68.

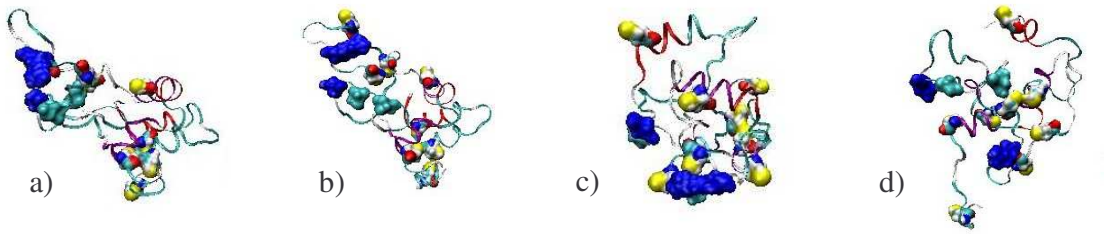


Figure 3.15: Graphical representations of unfolding at 500K, snapshots at a) 40ps b) 140ps c) 2ns and d) 3.5ns. Surfaces of residues 48 and 68 are colored blue and surfaces of residues 50 and 60 are colored light blue. Residues colored white, red and yellow are cysteine.

Residues 48 and 68 also appeared to be active in the unfolding simulations. With the added influence of 2 neighboring disulphide bonds the group can be construed to function as a gate, enabling unfolding. After the residues become fully detached from each other, serine residues at positions 50 and 60, which were stable and maintained their hydrogen bond during lower temperature simulations, started to fluctuate stronger, eventually dissociating.

4. NMR EXPERIMENTS OF HEW LYSOZYME

^{1}H and ^{13}C NMR experiments were carried out under different temperatures, pH values and in different chemical environments. Experiments were started with ^{1}H spectrums. After succeeding to reduce the water peak by employing solvent suppression methods, temperature array experiments were conducted. Water suppression definitely helped to observe chemical shifts of amino acid groups. Lysozyme was typically heated to 80˚C starting from 25˚C, at 5˚C increments. Free induction decays (FID) of excited nuclei were collected and Fourier transform was applied to obtain frequency domain data from the time domain data, namely, the FID. Each spectrum was manipulated to restore the phase information. Once the spectra were corrected, the effects of temperature or pH on different groups could be conveniently assessed. Kinetic analyses were achieved using built-in modules of VNMR software.

The above procedure was applied primarily to hydrogen nuclei to note any structural effects in varying the pH values between 2 and 11, and the effect of chemicals such as ethanol, methanol, glucose, acetone and thiamine.

^{13}C Natural abundance experiments were also conducted. Since ^{13}C is scarce in nature, the amount of protein sample used was increased 4 folds. Full decoupling and Nuclear Overhauser Effect methods were applied to encourage sharp spectra and strong signals, respectively.

Relaxation experiments were performed upon isotopically labeled methylated lysozyme. In particular a nonaqueous chemical modification procedure for trimethylating amino groups in proteins with iodomethane was followed [Taralp A., and Kaplan H., 1997]. Trimethylation of amino groups proceeds with retention of structure and bioactivity under appropriate reaction conditions. Nonaqueous protein methylations were carried out for up to 36h at 70 ˚C. The strategy of labeling protein with trimethylammonium groups not

only permits the study of chemical shifts of nuclei but also permits the study of protein dynamics. To prepare each sample, *in vacuo* methylation was carried out in a pyrex hydrolysis tube. Lysozyme was lyophilized from a solution of 50mM sodium phosphate buffer at pH 7.0. The neck of the tube was narrowed by narrowed by flame. ^{13}C-iodomethane was delivered to the bottom of the vessel using a micropipette and frozen by immersing the bottom half of the vessel in liquid nitrogen. A vacuum hose was applied to the opening and the tube was sealed under vacuum using s flame.

4.1 Materials

The VARIAN Inova 500MHz NMR system was used. ^{1}H Samples were analyzed with the Varian Indirect probe and ^{13}C samples were analyzed with the Varian Switchable probe. Pure hen egg-white lysozyme, deuterated water and acetonitrile solution were purchased from a commercial supplier (SIGMA).

4.2 Conditions

Protein folding rates are sensitive to a variety of environmental conditions, including temperature, pH, buffer, ionic strength, and the concentration and nature of any residual denaturant. At present, there is no standard set of conditions under which folding kinetic studies should be performed [Karen *et al.*, 2005].

In a typical session, the aqueous environment surrounding lysozyme was incrementally altered by introducing additives such as ethanol, methanol, sugar derivatives, formaldehyde and vitamins such as B_1, affording a series of related samples. Subsequently, sample were obtained at different pH values and temperatures, and the effects of environment on spectral profiles were related to protein structure.

4.2.1 Solvent System

A 90% H_2O – 10% D_2O solution was typically used. A high amount of H_2O was necessary to mimic cellular conditions, whereas D_2O was used to maintain the signal lock of the NMR instrument and to permit autoshimming.

4.2.2 Temperature

The experiments were typically begun at 25˚C. This temperature featured some technical advantages. It is slightly above room temperature and is readily maintained. Also, 25˚C has been the most commonly employed temperature in the literature [Karen *et al.*, 2005]. Thereafter, Proteins were heated to 80˚C at 5˚C increments. This procedure was adapted in the hope to gently unfold lysozyme, and to permit its refolding when temperature is decreased.

4.2.3 pH Value

Folding kinetics are sensitive to pH effects and thus it is preferable to adopt a consensus value for this parameter. The pH value of 7.0, which has been widely used in the literature, was used in these temperature experiments [Karen *et al.*, 2005].

4.2.4 Buffer

Due to its appropriate pKa value and its applicability to variable temperature and NMR experiments, 50 mM phosphate was used as buffer at pH 7. For similar reasons, borate buffer was used in the higher pH experiments. Several other buffers such as tris could have been used [Karen *et al.*, 2005].

4.3 Results

4.3.1 ^{1}H Spectra Analysis at Different Temperatures

^{1}H spectral analysis of proteins is a simple but powerful method, as it takes only minutes to get a clear spectrum that is unique for each macromolecule. Functional groups of the amino acids have chemical shift values, occupying regions of the spectrum. When conformational changes occur, the chemical environment of the nuclei naturally changes and the effects on chemical shift values may be observed in the ^{1}H spectra.

Without solvent suppression, groups that appeared near the DHO frequency were not observable. Their concealment made it difficult to deduce any dynamic changes as a whole. In figure 4.1 the water peak (DHO) may be easily seen at 4.8 ppm.

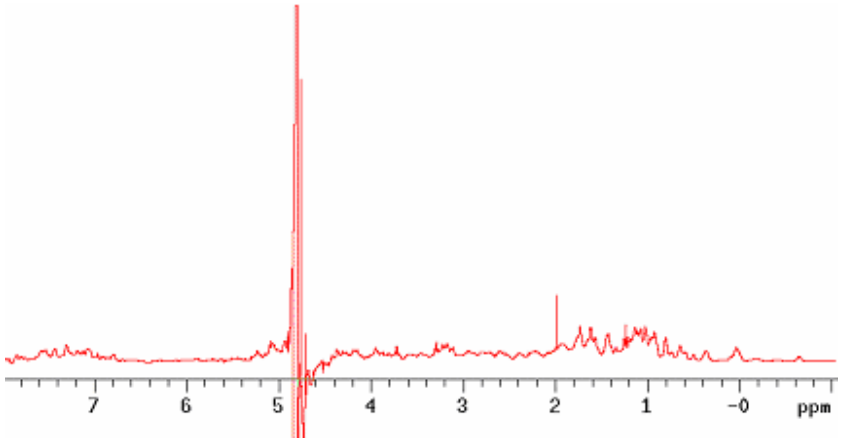


Figure 4.1: Lysozyme ^{1}H spectrum without water suppression reveals a DHO peak around 4.8 ppm. 256 scans. This strong signal reduced the quality of the whole spectrum.

After water suppression techniques were applied, the spectrum was greatly clarified, with the different resonances of the protein being visible as shown in figure 4.2.

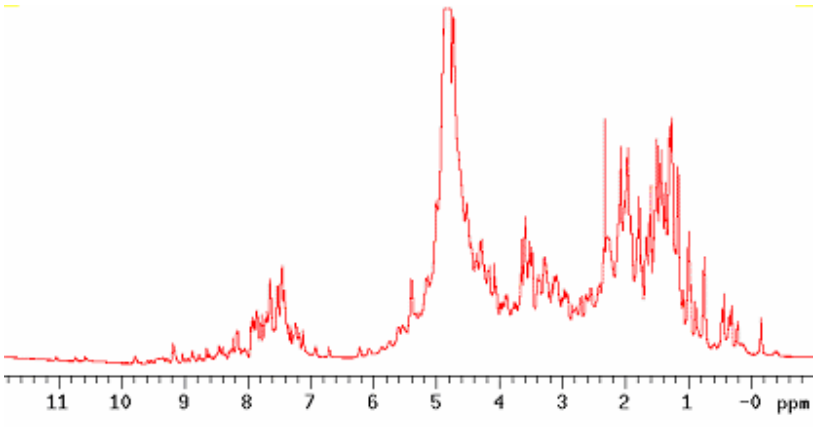


Figure 4.2: Lysozyme ^{1}H spectrum with water suppression. 256 scans. A built-in macro in the Varian software that employed presaturation suppressed the frequency of the solvent protons. As shown, signal intensity and resolution are now adequate to provide useful information.

4.3.2 Kinetic Experiments Under Different pH Values

One of the most important environmental factors determining protein functionality and stability is the pH value. The effect of pH has been well studied with different proteins in relation to understanding the folding mechanism. To reflect on different inter and intra-cellular environments, acidic and basic *in vitro* conditions were applied.

4.3.2.1 Acidic

The pH of the environment is an important factor in determining protein structure. Lysozyme is biologically active at slightly acidic pH values. Although it is not functional at very acidic conditions, it preserves its structural stability at low pH values.

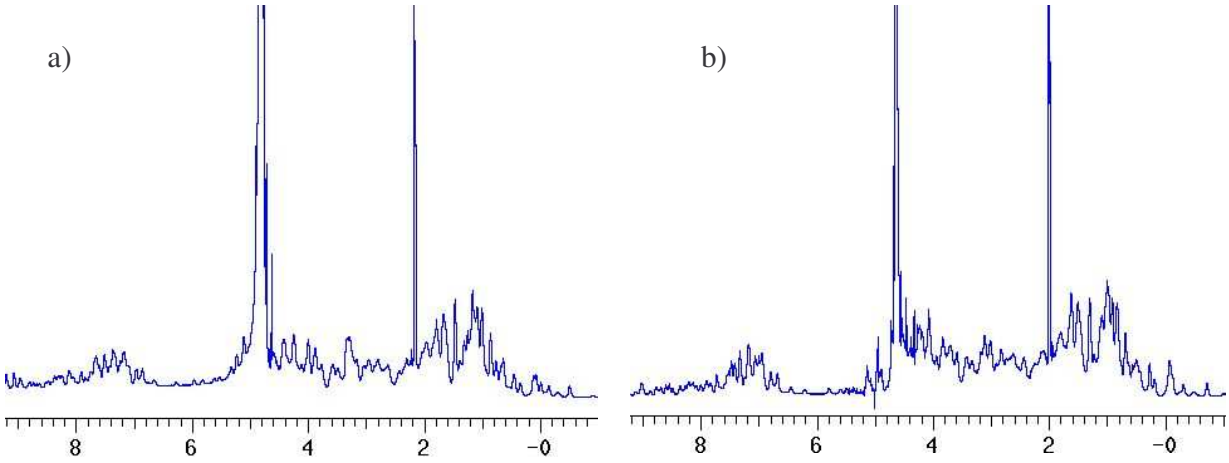


Figure 4.3: pH effect on protein structure at 35 ˚C, a) pH 2, and b) pH 4. 256 scans. Spectral profiles appear very similar

Figures 4.3 – 4.5 show the effect of pH on protein structure at different temperatures.

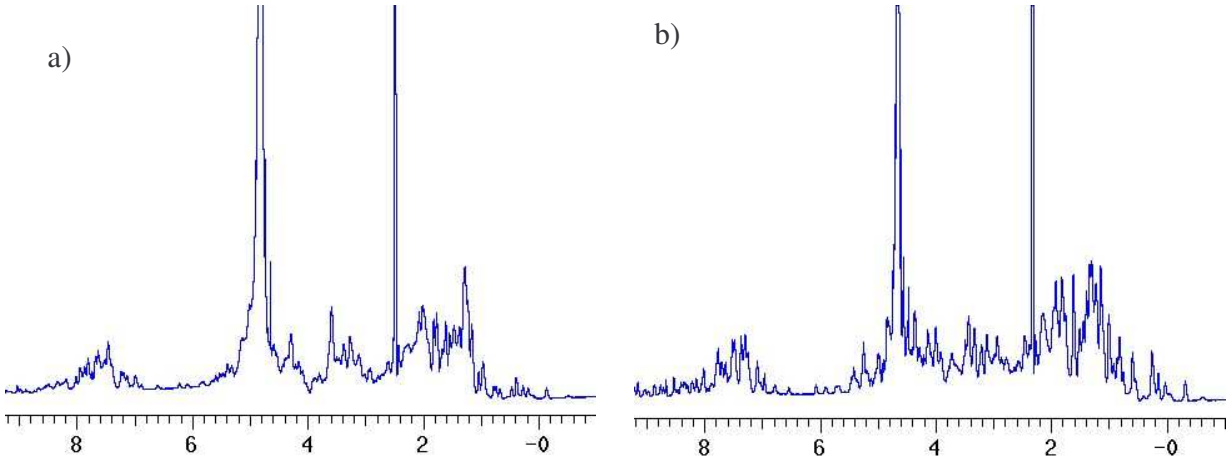


Figure 4.4: pH effect on protein structure at 65 C˚ a)pH 2 b) pH 4. 256 scans. Elevated temperature appears to affect only lysozyme at pH2.

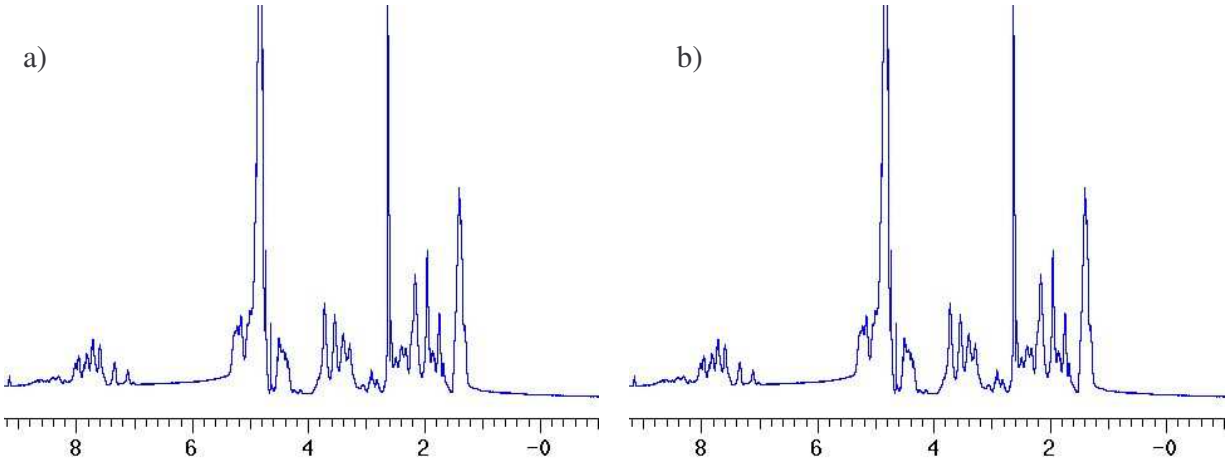


Figure 4.5: pH effect on protein structure at 80 ˚C a) pH 2 b) pH 4. 256 scans.

As temperature increased, lines were expected to sharpen since viscosity broadening becomes less significant at higher temperatures. Here, in contrast lines become broader as temperature was increased. It was deduced that lysozyme is not stable under these conditions. Structural changes and possibly a pre-gelation state were implied.

4.3.2.2 Basic

Several different conditions were applied to observe the effect of alkalinity on protein structure. ^{1}H spectra at pH 9 at different temperatures are shown in figure 4.6. Like most proteins, lysozyme is quite unstable at higher pH values. Temperature promoted aggregation or gelation at higher pH values was observed, vis-à-vis, the marked contrast at 80˚C (figure 4.6c).

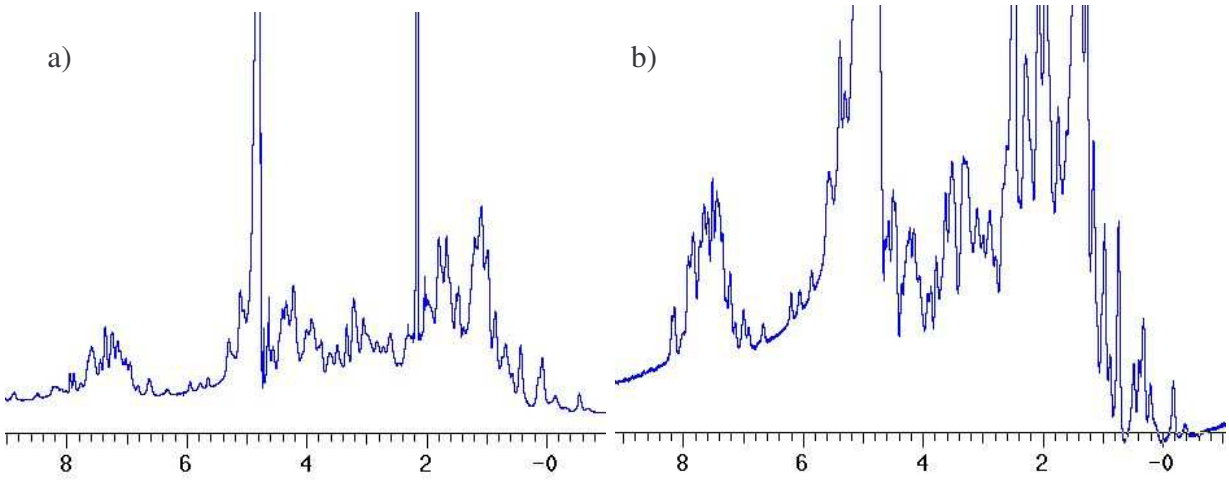


Figure 4.6: ^{1}H protein spectrum at pH 9 a) 35 ˚C b) 65 ˚C. 256 scans.

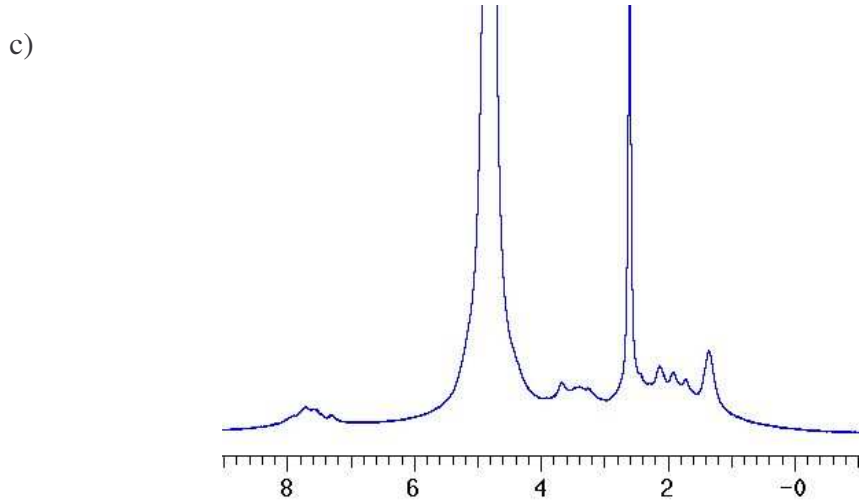


Figure 4.6c: ^{1}H spectrum of lysozyme at pH 9 for 80 ˚C. 256 scans.

Beyond pH 9.5, protein gelled easily in a short time (data not shown). When heated incrementally, protein appeared to begin unfolding at around 65˚C (spectral temperature array, data not shown).

4.4 Thiamine Effect on Protein Stability

Thiamine (Vitamin B_1) was the first vitamin identified in 1926 by Jansen and Donath while working on the antiberiberi factor from rice bran extracts. Thiamine mainly acts in α-ketoacid decarboxylation, in transketolation, and possibly in nerve conduction. It is phosphorylated when it crosses the intestinal epithelium, but enters the blood principally as free thiamine and diffuses down a concentration gradient in the liver, heart, kidneys, and brain. Nowadays thiamine is mainly detected as a consequence of extreme alcoholism and its manifestation is known as Wernicke-Korsakoff syndrome.

In the variable pH and temperature experiments, the effect of thiamine and other chemicals on proteins were examined. Lysozyme, wich gelated rapidly in basic pH and high temperature values, did not gelate at all in the presence of thiamine, presumably through its interaction with protein structure. Thiamine's stability has been highly studied, and ironically it decomposes when heated in solution. However its effect on protein structure has not been studied before and seemed beneficial in the case of lysozyme.

The spectra below were acquired at pH 9, using 50mg lysozyme and 10mg thiamine in D_20. Each spectrum was scanned 256 times except those following heat treatment, which was scanned 16000.

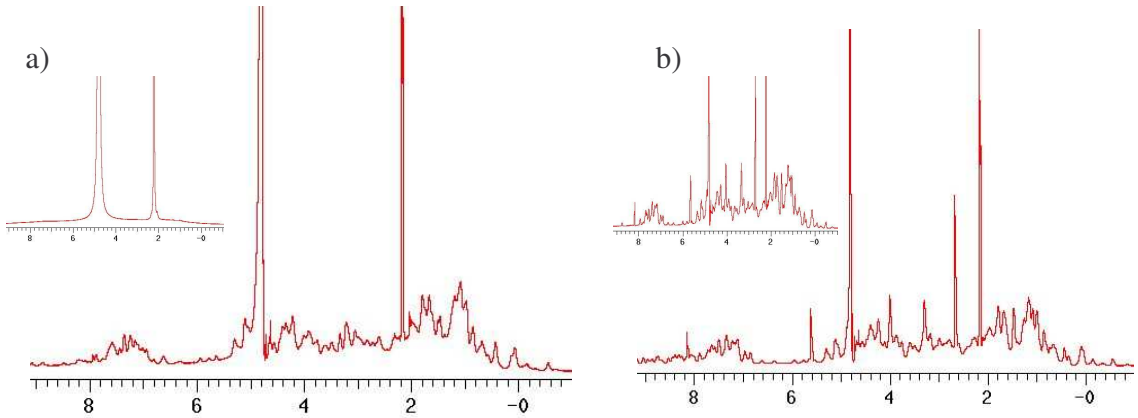


Figure 4.8: Thiamine effect on lysozyme stability at 35˚C and pH 9 a) Lysozyme without thiamine, b) with thiamine (insets shows spectra upon cooling of sample from 80˚C). 256 scans.

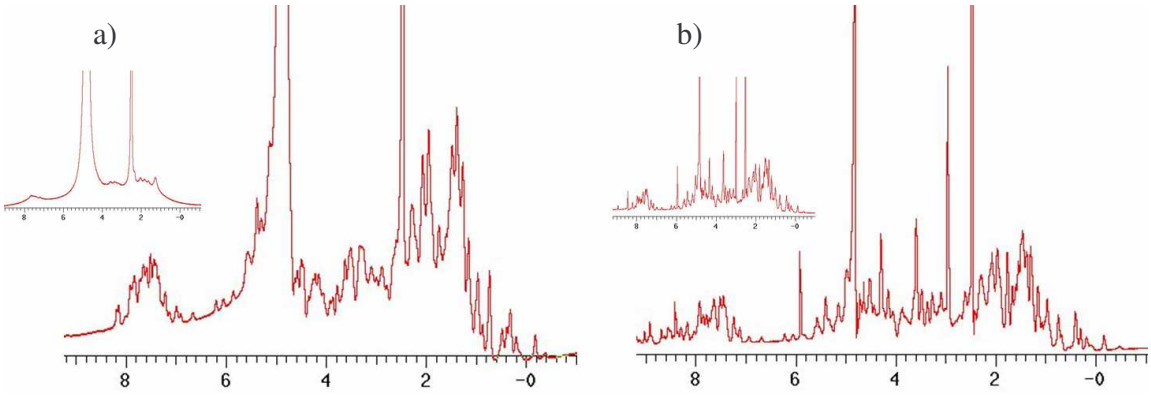


Figure 4.9: Thiamine effect on lysozyme stability at 65˚C pH 9 a) without thiamine b) with thiamine (insets shows spectra on cooling down from 80˚C). 256 scans.

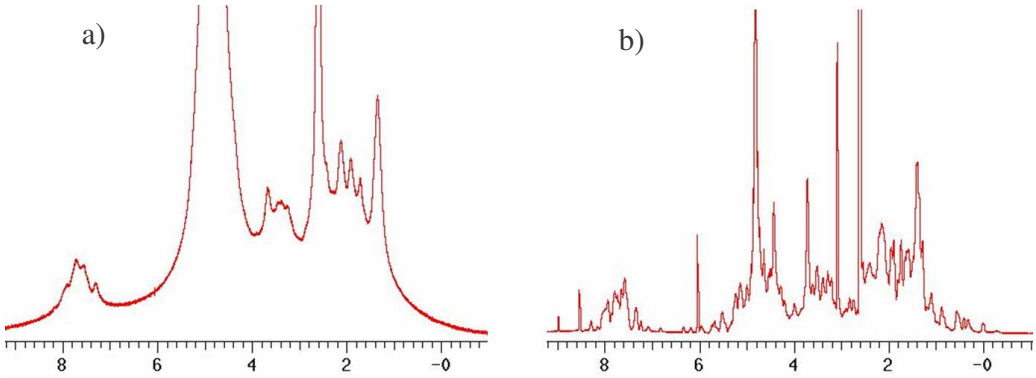


Figure 4.10: Lysozyme at 80˚C, pH 9 a) without thiamine b) with thiamine. 256 scans.

Figures 4.8, 4.9 and 4.10 illustrate that thiamine protected lysozyme against gelation at high pH and temperature values. Without thiamine, lysozyme quickly aggregated and gelated.

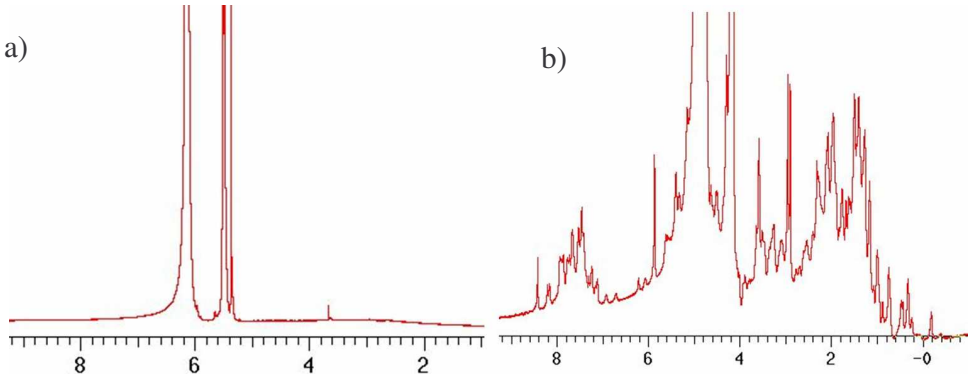


Figure 4.11: Heat treatment of lysozyme solution. After 10 hours at 70˚C, pH 9 a) Lysozyme without additive b) Lysozyme with Thiamine. 16000 scans. Permanent line broadening of proton spectrum in a, and apparent stability in b, clearly shows the suppression of the aggregation of denatured protein by Thiamine.

Since lysozyme starts unfolding approximately after 65˚C, it was deliberately kept at 70˚C for 10 hours with and without thiamine. Surprisingly, the protein spectrum with thiamine implied a relatively intact structure. While the solution without thiamine was gelated, the solution with thiamine was clean, as shown in figure 4.12.Indeed, the images below vividly illustrate the protective effect of thiamine.

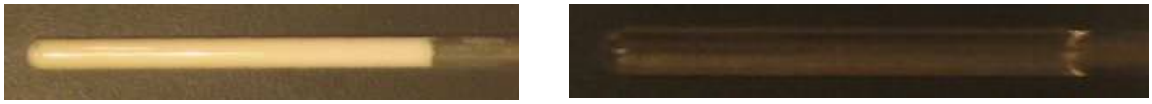


Figure 4.12: NMR tubes after heat treatment. Left, solution without thiamine; right, with thiamine.

4.5 Natural Abundance ^{13}C Experiments

^{13}C exists naturally in very small proportions, so NMR experiments of ^{13}C without isotopic enrichment can be inconvenient to observe. In figure 4.13 & 4.14 200mg of lysozyme was dissolved in 700µl D_2O solution. With full decoupling and full NOE enhancement applied, the ^{13}C nuclei chemical shifts could barely be observed after 1024 scans at temperatures ranging between 25˚C to 80˚C.

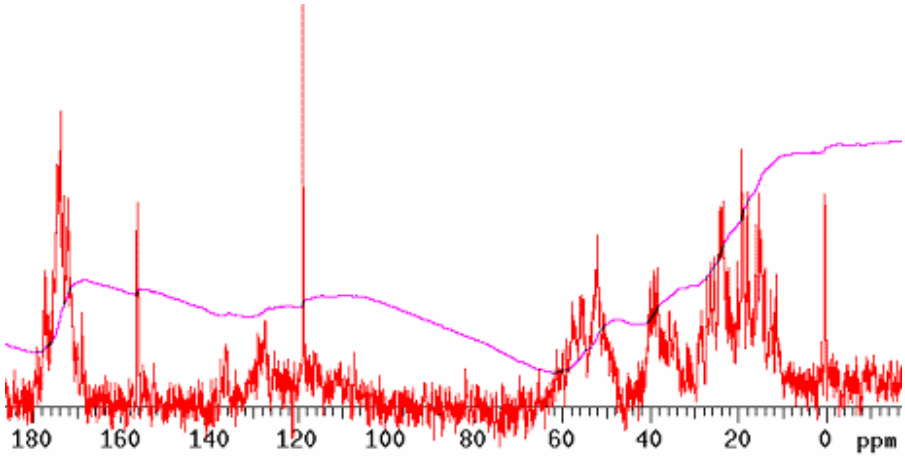


Figure 4.13: Natural abundance ^{13}C spectrum, at room temperature. 1024 scans.

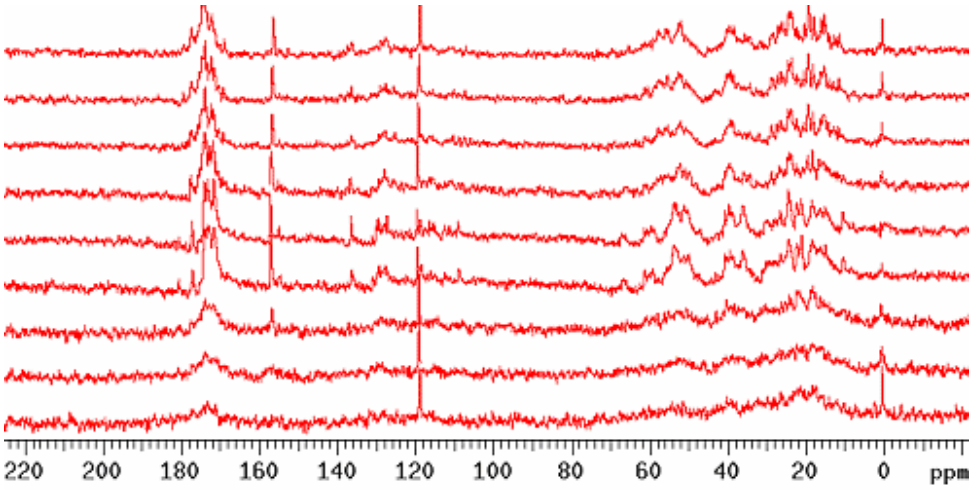


Figure 4.14: Natural abundance lysozyme ^{13}C spectra swept at different temperatures, (from top to bottom) 25, 40, 45, 65, 80, 65, 45, 40, 25˚C. 1024 scans.

As it is clearly implied in figure 4.14, lysozyme did not refold properly in natural abundance ^{13}C experiments. It was assumed that when the solution was cooled down, the lysozyme molecules associated with each other, certainly preventing appropriate refolding. Although correct refolding was implied in the thiamine-protected ^{1}H experiments, no thiamine was used in the carbon experiments. This contrast was a good indication of the protective effect of thiamine on lysozyme structure. Presumably, thiamine somehow created a layer around the proteins, preventing their bonding to each other.

4.6 ^{13}C Methylation Experiments

NMR relaxation experiments of isotopically labeled proteins provide a wealth of information on reorientational global and local dynamics on nanosecond and subnanosecond timescales for folded and unfolded proteins in solution. Recent methodological advances in the interpretation of relaxation data have led to a better understanding of the overall tumbling behavior, the separability of internal and overall motions, and the presence of correlated dynamics between different nuclear sites, as well as new insights into the relationship between reorientational dynamics related to primary and tertiary protein structure. Some of the newer methods are particularly useful when dealing with nonfolded protein states [Bruschweiler, 2003].

Here 20mg of lysozyme was methylated using the *in vacuo* labeling method of Taralp [Taralp, A., and Kaplan H.,1997] primarily at His and Lys side-chains along the surface. ^{13}C-labelled iodomethane could potentially provide useful information to probe the overall or local surface dynamics of proteins at residues with reactive side chains [Taralp, A., and Kaplan H.,1997].

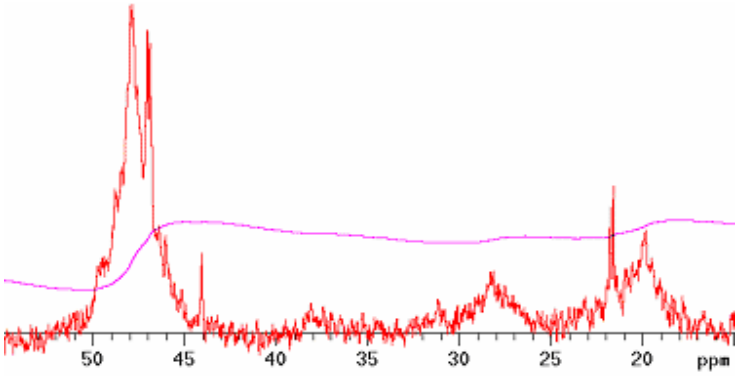


Figure 4.15: ^{13}C methylated lysozyme, 700µl D_2O, 5µl acetonitrile, 16000 scans, 25˚C.

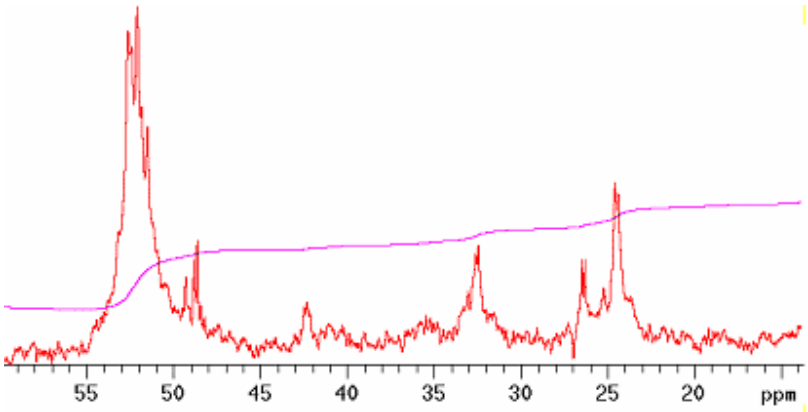


Figure 4.16: ^{13}C methylated lysozyme, 700µl D_2O, 5µl acetonitrile, 16000 scans, 60˚C.

4.7 Relaxation Experiments

Backbone and methyl side chain NMR relaxation measurements for several proteins are beginning to reveal the role of protein dynamics in protein stability and ligand binding [Spyracopoulos, 2005].

Since the ^{13}C T_1 values are sensitive monitors of the rates and amplitudes of the internal motions of the protein, a comparison was made to test the extent to which MD

provides an accurate depiction of the internal motions of proteins [Balasubramanian *et al.*, 1994].

^{13}C relaxation parameters were measured for ^{13}C methyl-lysozyme at room temperature. Experiments were repeated with the same parameters at 35˚C and 65˚C. T_1 and T_2 values were calculated for 7 different tau values (0.875s, 1.75s, 3.5s, 7s, 14s, 28s, and 56s) at 25˚C. 20mg labeled lysozyme was used in all relaxation experiments. The average T_1 value was 13.79 and average T_2 was 14.2.

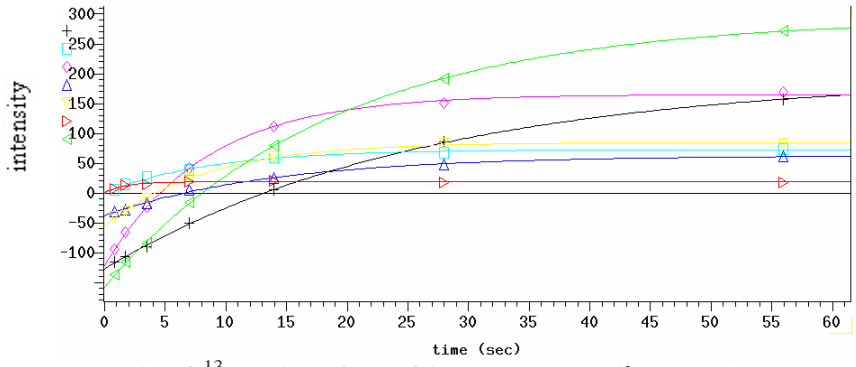


Figure 4.17: T_1 graph of ^{13}C relaxation of lysozyme at 25˚C. 1024 scans. Each solid line marks a different peak in the spectrum.

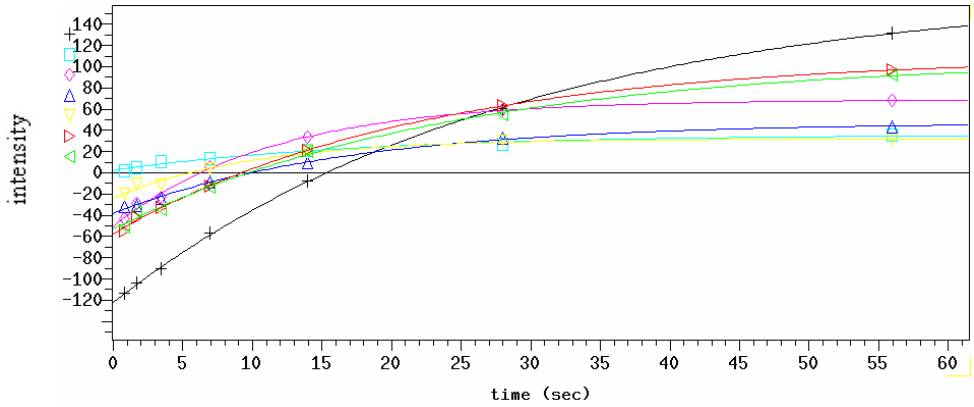


Figure 4.18: T_1 graph of ^{13}C relaxation of lysozyme at 35˚C. 1024 scans.

At 35˚C, the average T_1 value was 19.02 and average T_2 value was 19.25. A relatively high increase in relaxation time was noted as hydration and viscosity effects were presumably at work at this temperature.

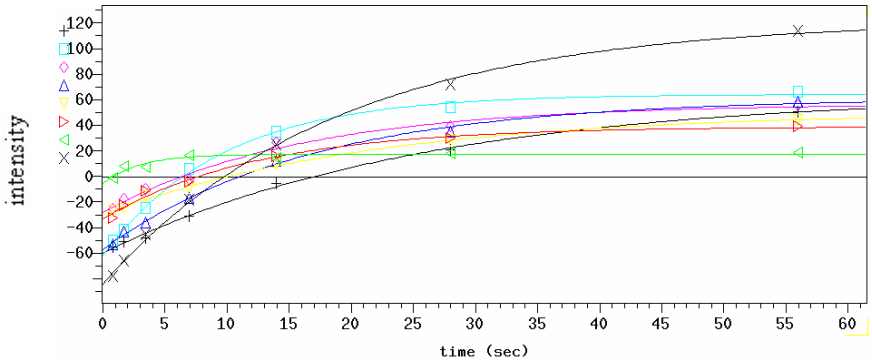


Figure 4.19: T_1 graph of ^{13}C relaxation of lysozyme at 65˚C. 1024 scans.

At 65˚C, the average T_1 value was 22.4 and average T_2 value was 16.6. It was assumed that after approximately after 65˚C, the protein starts unfolding, leading to a more crowded environment. Although not directly applicable, the results seem to be related to computational analysis of relaxation at different temperatures, wherein relaxation during the first 20ps was larger in simulations depicting examples of unfolding.

5. CONCLUSIONS & FUTURE WORK

"The important thing is not to stop questioning. Curiosity has its own reason for existing..."
Albert Einstein

The fundamental events of proteins are very difficult to study, even *in vitro*. These difficulties formed a rational basis for experimentalist and theoretician to work together, particularly as methods related to both areas have developed rapidly and given many satisfactory results. Also, the synergy of working together has increased the understanding of protein folding and cellular mechanisms. In this study, a combined NMR and MD approach permitted a comparison of results from more than one viewpoint.

In the first part, many simulations of lysozyme were designed and carried out. Heat capacity, backbone fluctuation, relaxation, radius of gyration, water accessible surface area and temperature autocorrelation functions were calculated. Heat capacity results pointed out that at the unfolding temperature, heat capacity values tend to drop as sharp conformational changes occur, and energy is spent to break non-covalent bonds. After protein fully unfolds protein is quite free to move, resulting in a maximum heat capacity value. Then a possible aggregation of protein drops heat capacity to a medium value. When results with the DSC experiments were compared, it was observed that heat capacity calculations were in good agreement with the experimental data until just above physiological values. A second sharp peak in the DSC data may be explained by a local conformational change, causing a reformation of an important bond in lysozyme, which is broken in a short time (data not shown).

Analysis of simulation results showed some interesting similar patterns, which could potentially be fingerprints of unfolding proteins and related to distinct intermediate structures. For this purpose, simulations should be revised and repeated with different

proteins and extent the analysis methods to get a better understanding.

Many experiments with NMR were designed and analyzed to study the effects of chemicals and environment on unfolding and stability of the protein. Kinetic analysis showed that lysozyme had a energy bottleneck of unfolding around 65-70 ˚C. It unfolded after 70˚C in light of gelation data and refolded when it was cooled, though the reversibility of unfolding was not always guaranteed and depended on the environment.

It was observed that vitamin B_1, whose deficiency causes neurodegenerative illnesses in humans, apparently protected protein structure at high pH values and temperatures. Although thiamine is well studied and is an important food supplement, its stabilizing effect on protein structure had not been studied before. While the mode of action of vitamin B_1 remains to be elucidated, it would follow to reason that it is related to a direct physico-chemical interaction between protein surface groups and functional group moieties of the vitamin.

To summarize, the major findings of this study were;

- Heat capacity changes between 150K and 320K are similar to calculations previously carried out for BPTI by Baysal and Atilgan (figure 3.3).
- The heat capacity of protein folding at 70˚C is maximum for lysozyme.
- The radius of gyration spectra, which represents compactness of proteins, was similar to previous studies (figure 3.5).
- The radius of gyration spectra at 500K was similar to the spectra of mean values at each temperature, showing that increasing the simulation temperature does change the intermediate structures but shortens the simulation time (figure 3.6).
- The change of water accessible surface area correlated with the radius of gyration data at 500K, showing solution effect on protein structure (figure 3.7).
- Cα fluctuations were maximum around the neighbors of proline 70, namely, arginine, threonine, glycine and serine residues. In those residues, the hydrogen bond of arginine 68 & aspartic acid 48 and bond of serine 50 & serine 60 were found to be very important in maintaining native protein structure (figures 3.13, 3.14).
- Including thiamine at basic conditions and high temperatures appears to stabilize protein structure (figures 4.8, 4.9, 4.10, 4.11).

The overall results of experiments and findings supported the notion that proteins follow a certain route for unfolding.It follows that observing the mechanism of unfolding and refolding may help to better understand protein folding problem. Also, the stability of many proteins can be controlled, at least in part, by appropriate use of additives apart from conventional ones such as sugars.

As a future work, simulations with thiamine and lysozyme together are planned at high temperatures. To generalize the notion of the existence of a fingerprint to unfolding, simulations are to be repeated with different proteins such as BPTI. 2D NMR experiments of lysozyme will be conducted to strengthen findings related to the thiamine effect. Also, other water soluble vitamins will be tested to ascertain if there is a general protective effect of vitamins on protein structure.

6. REFERENCES

[1]Allen, M. P., D. J. Tildesley. 1989. *"Computer Simulation of Liquids",* Clarendon Press.

[2]Andersen, H. C. 1983. *Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations*. J.Comput. Phys. 52:24-34.

[3]Arnold, J. T., S. S. Dharmatti, M. M. Packard. 1951. *Chemical effects on nuclear induction signals from organic compounds*. Chem. Phys. 19: 507.

[4]Bahar, I., B. Erman, T. Haliloglu, R. L. Jernigan. 1997. *Efficient characterization of collective motions and interresidue correlations in proteins by low-resolution simulations.*
Biochemistry. Nov 4;36(44):13512-23. Review.

[5]Baianu, I. C., T. F. Kumosinski, P. J. Bechtel, A. Mora, L. T. Kakalis, P. Yakubu, P. Myers-Betts, T. C. Wei. 1991. *Molecular dynamics of water in foods and related model systems: multinuclear spin relaxation studies and comparison with theoretical calculations.,* Adv Exp Med Biol. 302:517-40.

[6]Balasubramanian, S., R. Nirmala, D. L. Beveridge, P. H. Bolton. 1994. *Comparison of the ^{13}C relaxation times and proton scalar couplings of BPTI with values predicted by molecular dynamics.* J Magn Reson B. Jul;104(3):240-9.

[7]Baysal, C., A. R. Atilgan. 2002. *Relaxation kinetics and the glassiness of proteins: the case of bovine pancreatic trypsin inhibitor.* Biophys J. Aug;83(2):699-705.

[8]Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. 2000. *The Protein Data Bank.* Nucleic Acids Research. **28** pp. 235-242.

[9]Blake, C.C.F. *et al.* 1965. *The structure of hen egg white lysozyme: a three dimensional Fourier synthesis at 2 Å resolution.* Nature 196:1173−1176.

[10]Bloch, F., W. W. Hansen, M. Packard. 1946. *The nuclear induction experiment.* Phys. Rev. 70: 474-485.

[11]Brooks, R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. *CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations*, J. Comp. Chem. 4:187-217.

[12]Bruschweiler, R. 2003. *New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins*., Curr. Opin. Struct. Biol. Apr;13(2):175-83.

[13]Canfield, R. 1963. *The Amino Acid Sequence of Egg White Lysozyme*. J. Biol. Chem. 238:2698.

[14]Case, D. A. 1998. *The use of chemical shifts and their anisotropies in biomolecular structure determination,* Curr. Opin. Struct. Biol. Oct;8(5):624-30.

[15]Case, D. A. 2002. *Molecular dynamics and NMR spin relaxation in proteins.* Acc. Chem. Res. Jun;35(6):325-31.

[16]Clark, R., J. P. Waltho. 1997. *Protein Folding Pathways and Intermediates,* Curr. Opin. Struct Biol. 8:400-410.

[17]Cooper, A. 2000. *Heat Capacity of hydrogen-bonded networks: an alternative view of protein folding thermodynamics*. Biophysical Chemistry. 85:25-39.

[18]Cramer, J. A., V. T. Marchesi, I. M. Armitage. 1980. *Reversible trifluoroacetic acid-induced conformational changes in glycophorin A as detected by proton nuclear magnetic resonance spectroscop.* Biochim. Biophys. Acta. Jan;25;595(2):235-43.

[19]Cross, T. A., F. Tian, M. Cotten, J. Wang, F. Kovacs, R. Fu. 1999. *Correlations of structure, dynamics and function in the gramicidin channel by solid-state NMR spectroscopy.* Novartis Found. Symp. 225:4-16; discussion 16-22.

[20]Daggett, V., P. A. Kollman, I. D. Kuntz. 1991. *A molecular dynamics simulation of polyalanine: an analysis of equilibrium motions and helix-coil transitions.* Biopolymers. Aug;31(9):1115-34.

[21]Deschenes, L. A., D. A. Vanden Bout. 2001. *Single-molecule studies of heterogeneous dynamics in polymer melts near the glass transition*. Science. Apr 13;292(5515):255-

8.
[22]Diamond, R. 1974. *Real-space refinement of the structure of hen egg-white lysozyme*. J.Mol.Biol. 82:371.

[23]Elwell, M., J. Schellman. 1975. *Phage T4 lysozyme. Physical properties and reversible unfolding.*, Biochim. Biophys. Acta. Mar 28;386(1):309-23.

[24]Fersht, A. R., V. Daggett. 2002. *Protein folding and unfolding at atomic resolution.* Cell. Feb 22;108(4):573-82.

[25]Fersht, A. R., V. Daggett. 2003. *Is there a unifying mechanism for protein folding?*. Trends. Biochem. Sci. Jan;28(1):18-25.

[26]Fleming, A. *On a remarkable bacteriolytic element found in tissues and secretions.* Proc. Roy. Soc. Ser. B. 93:306-17.

[27]Fraunfelder, H., McMahon B. *Dynamics and function of proteins: The search for general concepts*. Proc. Natl. Acad. Sci. U.S.A 95:4795-4797.

[28]Gorter, C. J. 1936. "*Negative result of an attempt to detect nuclear magnetic spins*," Physica 3:995-998.

[29]Haile, J. M. 1992. *Molecular Dynamic Simulations: Elementary Methods*, John Wiley & Sons, New York.

[30]Hirai, M., M. Koizumi, T. Hayakawa, H. Takahashi, S. Abe, H. Hirai, K. Miura, and K. Inoue. 2004. *Hierarchical Map of Protein Unfolding and Refolding at Thermal Equilibrium Revealed by Wide-Angle X-ray Scattering*. Biochemistry. 43:9036-9049.

[31]Howarth, O. W. 1979. *The thermal unfolding of ribonuclease A. A 13C NMR study*. Biochim Biophys. Acta. Jan 25;576(1):163-75.

[32]Humphrey, W., A. Dalke, and K. Schulten. 1996. "*VMD - Visual Molecular Dynamics"*, J. Molec. Graphics. 14:33-38.

[33]Johnson, L. N. 1998. *The Early history of Lysozyme*. Nature Structural Biology. 5:942–944.

[34]Jolles, J., P. Jolles. 1961. *Chemical structure of the lysozyme of chicken egg white:the developed formula*. C. R. Hebd Seances Acad. Sci. Dec 4;253:2773-5.

[35]Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999 NAMD2: *Greater scalability for parallel molecular dynamics***.** Journal of Computational Physics**,** 151:283-312.

[36]Karplus, M., G. A. Petsko. 1990. *Molecular dynamics simulations in biology.* Nature. Oct 18;347(6294):631-9.

[37]Kazmirski, S. L., V. J. Daggett. 1998. *Non-native Interactions in Protein Folding Intermediates: Molecular Dynamics Simulations of Hen Lysozyme*, J. Mol. Biol. 284:793-806.

[38]Kuwajima, K., F. X. Schmid. 1984. *Experimental studies of folding kinetics and structural dynamics of small proteins*. Advances in Biophysics. 18:43-74.

[39]Leach, A. R. 1996. *Molecular Modelling: Principles and Applications*. Addison Wesley Longman: Singapore.

[40]Levinthal, J. 1968. *Are Their Pathways for Protein Folding*. J. Chem. Phys. 65:44-45.

[41]Lewitt, M. 1983. *Molecular dynamics of native protein. II. Analysis and nature of motion.* J. Mol. Biol. Aug 15;168(3):621-57.

[42]Louise-May, S., P. Auffinger, E. Westhof. 1996. *Calculations of nucleic acid conformations.* Curr. Opin. Struct. Biol. Jun;6(3):289-98.

[43]MacKerell, Jr. A.D. 2001. *"Atomistic Models and Force Fields," in Computational biochemistry and biophysics.* O.M. Becker Ed. 7-38.

[44]Mark, E., W. F. van Gunsteren. 1992. *Simulation of the Denaturation of Hen Egg White Lysozyme: Trapping the Molten Globule State*. Biochemistry. 31;7745-7748.

[45]Matagne, S. E. Radford, C. M. Dobson. 1997. *Fast and Slow Tracks in Lysozyme Folding: Insight into the Role of Domains in the Folding Process.* J. Mol. Biol. 267:1068-1074.

[46]Maxwell, K L., D. Wildes, A. Zarrine-Afsar, M. A. De Los Rios, A. G. Brown, C. T. Friel, L. Hedberg, J. C. Horng, D. Bona, E. J. Miller, A. Vallée-Bélisle, E. R.G. Main, F. Bemporad, L. Qiu, K. Teilum, N. D. Vu, A. M. Edwards, I. Ruczinski, F. M. Poulsen, B. B. Kragelund, S. W. Michnick, F. Chiti, Y. Bai, S. J. Hagen, L. Serrano, M. Oliveberg, D. P. Raleigh, P. Wittung-Stafshede, S. E. Radford, S. E. Jackson, T. R. Sosnick, S. Marqusee, A. R. Davidson and K. W. Plaxco. 2005. *Protein folding:*

Defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins. Protein Science. 14:602-616.

[47]Mayor, U., C. M. Johnson, V. Daggett, A. R. Fersht. 2000. *Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation.* Proc. Natl. Acad. Sci. USA. Dec 5;97(25):13518-22. Erratum in: Proc Natl. Acad. Sci USA 2001 Jan 16;98(2):777.

[48]McCammon, J., B. R. Gelin, M. Karplus. 1997. *Nature*. *267*(5612), 585–90.

[49]Motoshima, H., S. Mine, K. Masumoto, Y. Abe, H. Iwashita, Y. Hashimoto, Y. Chijiiwa, T. Ueda, T. Imoto. 1997. *Analysis of the stabilization of hen lysozyme by helix macrodipole and charged side chain interaction*. J. Biochem. (Tokyo) 121:1076.

[50]Nozaka, M., K. Kuwajima, K. Nitta, S. Sugai. 1978. *Detection and characterization of the intermediate on the folding pathway of human alpha-lactalbumin.,* Biochemistry. Sep 5;17(18):3753-8.

[51]Pauling, L. Rb. Corey, Hr. Branson. 1951. *The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain*. Proc. Natl. Acad. Sci. USA. Apr;37(4):205-11.

[52]Prabhu, N. H., and Sharp, K.A. 2005., Heat capacity in proteins. Annu. Rev. Chem. 56:521-48

[53]Purcell, E. M., H. C. Torrey, R. V. Pound. 1946. *Resonance absorption by nuclear magnetic moments in a solid*. Phys. Rev. 69:378.

[54]Schlick, T. 2002. Molecular *Modeling and Simulation An interdisciplinary guide* , Springer. ISBN 0-387-95404-X USA.

[55]Sen, U. 2004. *Dynamic Simulations of Novel Oligomers containing fluorinated segments.* Thesis Study. Sabanci University.

[56]Shakhnovich, I.. 1979. *Theoretical Studies of Protein-Folding Thermoldynamics and Kinetics*, Curr. Opin. Struct. Biol. 7:29-40.

[57]Shimizu, H. 1979. *Dynamic cooperativity of molecular processes in active streaming, muscle contraction, and subcellular dynamics: the molecular mechanism of self-organization at the subcellular level*. Adv. Biophys. 13:195-278.

[58]Smith, T. 1999. *Prelude to NMR studies of protein folding*. Nat. Struct. Biol. Jul;6(7):608.
[59]Spyracopoulos, L. 2005. *Thermodynamic interpretation of protein dynamics from NMR relaxation measurements*. Protein Pept. Lett. Apr;12(3):235-40.

[60]Svergun, D. I., C. Barberato & M. H. J. Koch. 1995. CRYSOL Wintel/Unix/Linux version 2.5. Version (LMAX=50) for small and wide angles. *J. Appl. Cryst.*, 28:768-773.

[61]Svergun, D. I., M. H. Koch. 2002. *Advances in structure analysis using small-angle scattering in solution*, Current Opinion in Structural Biology. Oct 1;12(5):654-660.

[62]Taralp, A., and H. Kaplan. 1997. Journal of Protein Chemistry 16:183-193.

[63]van de Ven, F. J., C. W. Hilbers. 1987. *Reversible unfolding of ribosomal protein E-L30: an NMR study.*Biochemistry. Aug 25;26(17):5548-55.

[64]Wildegger, G., T. J. Kiefhaber. 1997. *Three State Model for Lysozyme Folding: Triangular Folding Mechanism With an Energetically Trapped Intermediate*, J. Mol. Biol. 270:294-304.

[65]Williams, M. A., J. M. Thornton, J. M. Goodfellow. 1997. *Modelling Protein Unfolding: Hen Egg-White Lysozyme*. Protein Eng. 10,895-903.

[66]Zaccai, G. 2000. *How soft is a protein? A protein dynamics force constant measured by neutron scattering.* Science. 288:1604-1607.