UNDERSTANDING PROTEIN DYNAMICAL TRANSITION AND PROTEIN-WATER INTERACTIONS FROM DIELECTRIC RELAXATION CALCULATIONS

by

Mehmet Irmak Sirer

Submitted to the Graduate School of Sabancı University in partial fulfuillment of the requirements for the degree of Master of Science

Sabancı University

Spring 2006

© Mehmet Irmak Sirer 2006

All Rights Reserved

ABSTRACT

UNDERSTANDING PROTEIN DYNAMICAL TRANSITION AND PROTEIN-WATER INTERACTIONS FROM DIELECTRIC RELAXATION CALCULATIONS

Dielectric properties of an aqueous lysozyme solution were calculated from 2 ns long MD simulations in the temperature range of 150-300 K and an 4 ns long simulation at 300 K. Static and frequency dependent dielectric constants of the system were calculated from auto- and cross-correlations of its three components (protein, water, ions). Cole-Cole plots for protein, water and the total solution were obtained. Emergence of an intense protein-water interaction above the dynamical transition between 190 K and 210 K was evidenced by the presence of protein effects in the water components of the Cole-Cole plots and frequency dependent dielectric constants at and above 210 K. Backbone and side chain torsion angle trajectories for surface loop residues within this range of temperatures were calculated. Also, water molecules around side chains were labeled and monitored individually, and radial distribution functions of water around the side chains and in the bulk water were obtained. These data were used to support a model that accounts for the interaction between surface water and protein components, resulting in high mobility of the side chains at the transition temperature range. The water molecules in the vicinity of the protein surface are then propelled into the bulk for a much different electrostatic effect than is immediately expected of the known properties of water alone. The functional protein, therefore, exists as an integral part of a larger protein-water system that cannot be decoupled. The water molecules may even be thought of as information carriers that make other nearby biological molecules aware of the presence of the protein.

ÖZET

DİELEKTRİK GEVŞEME HESAPLARINDAN PROTEİN DİNAMİK DEĞİŞİMİNİ VE PROTEİN-SU ETKLEŞİMLERİNİ ANLAMAK

Su içeren bir lizozom çözeltisinin dielektrik özellikleri, 150-300 K sıcaklık aralığı içinde 2 ns uzunluğunda ve 300 K için 4 ns uzunluğunda gerçekleştirilen moleküler dinamik simulasyonlarından hesaplandı. Sistemin statik ve frekansa bağlı dielektrik sabitleri, üç bileşenin (protein, su, iyonlar) kendileriyle ve birbirleriyle olan korelasyon fonksiyonlarından hesaplandı. Protein, su ve bütün çözelti için Cole-Cole grafikleri çizildi. 190 K – 210 K arasında gerçekleşen dinamik değişimden sonra kuvvetli bir protein-su etkileşimin başlangıcı, su için çizilen Cole-Cole grafiklerinde ve frekansa bağlı dielektrik sabitlerinde 210 K üzerinde görülen protein etkisiyle kanıtlandı. Yüzey aminoasitlerinin çatısal ve yan zincir dihedral açıları hesaplandı. Yan zincirlerin etrafındaki su molekülleri etiketlenip tek tek izlendi ve yüzeye yakın ve uzak suların radyal dağılım fonksiyonları hesaplandı. Bu verilere dayanarak yüzey suları ve protein etkileşimine dair bir model geliştirildi. Geçiş sıcaklığında bu etkileşimin yan zincirlere yüksek hareketlilik kazandırdığı, bu hareketin yüzey sularını dışarıya iterek suların rotasyonunu sağladığı ve böylece suyun tamamına beklenmedik bir elektrostatik etki yüklediği sonucuna varıldı. Buna göre işlevsel protein, daha büyük ve ayrılamaz bir protein-su sisteminin bir parçasıdır. Su molekülleri de etraftaki diğer biyolojik molekülleri proteinin varlığından haberdar eden bilgi taşıyıcaları olarak düşünülebilirler.

ACKNOWLEDGMENTS

I am deeply grateful to my advisor Canan Atılgan. I don't believe I deserve her understanding and support. I can only hope to make up for my shortcomings in the future by becoming half the productive and inspiring scientist she is. And if I manage to do that, it is only possible because of the model she put before me. My aim in life is to become a successful young teacher and scientist like her. And I don't think I would be exaggerating when I say it was her who saved my scientific career, since switching to computational, theoretical research brought the inspiration I desperately needed.

I would also like to thank Ali Rana Atılgan for his brightly burning eyes, shooting rays of intelligence. A raised eyebrow, followed by a first question about a result means you are in for a thrill ride! Questions follow, a chain of ideas springs to life, and you live the room ultimately motivated.

Deniz Turgut has been a great roommate, a provocative labmate and a trusted supporter. I could not imagine working and finishing this thesis without him. Whether we are discussing a book, a fragment or a subtlety in Galois theory, I always admire his intelligence. My friends, Kerem Gören, Eren Şimşek, Sinan Yördem, Osman Burak Okan, Burcu Köktürk, Işıl Nalbant, Pınar Önal and others, have been my fellow jedi knights. Without them, the dark Empire would have invaded this thesis.

I am most thankful to my parents and family. Their love is a safety net whenever I feel lost. I hope this work is worthy of them. A million thanks to the best parents in the world. My part in this work is dedicated to them.

I am also grateful to Tübitak for providing me with the National Graduate Scholarship. And finally I thank Ceyda Düvenci for her neverending support and beauty, which have propelled me, especially in the last phases of the study.

TABLE OF CONTENTS

ABSTRACT iv
ÖZET v
ACKNOWLEDGMENTS vi
TABLE OF CONTENTS vii
LIST OF FIGURES ix
LIST OF TABLES x
LIST OF SYMBOLS AND ABBREVIATIONS xi
1. INTRODUCTION 1
2. THEORETICAL BACKGROUND 10
2.1 Molecular Dynamics 10
2.2 General Theory of Dielectrics, Susceptibility and Permittivity 11
2.3 Dielectric Theories on Solutions of Macromolecules 14
2.3.1 Kirkwood-Fröhlich Teory 14
2.3.2 Direct Approach 15
2.3.3 Dielectric Field Equation (DFE) 16
2.3.4 Adjustment of Dielectric Boundary Conditions 17
2.3.5 Linear Response Theory 17
2.3.6 Handling Charged Proteins 19
2.3.7 Simulation Rules for Dielectric Calculations 21
2.3.8 Calculation of Susceptibility Using Computer Adaptive Linear Response Theory 22
3. SIMULATION DETAILS 25
4. RESULTS AND DISCUSSION 26
4.1 Dielectric Relaxation: Correlation Functions 26
4.1.1 Autocorrelations 26
4.1.2 Cross-Correlations 29
4.1.3 Double Exponential Fit 30
4.2 Dielectric Constants 33
4.2.1 Static Dielectric Constants 33
4.2.2 Frequency Dependent Complex Dielectric Constants 34

4.2.3 Cole-Cole plots....36
4.3 Side Chains and Water Molecules....37
4.3.1 Mobility of Side Chains....37
4.3.2 Radial Distribution Functions....39
5. CONCLUSIONS AND FUTURE WORK....43

LIST OF FIGURES

Figure 2.1 Response of a dielectric medium containing polar molecules to an applied electric field between parallel plates [41] 12

Figure 4.1 Protein component total dipole moment autocorrelation functions $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$ 27

Figure 4.2 Water component total dipole moment autocorrelation functions $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$ 27

Figure 4.3 Current autocorrelation function at 290 K 28

Figure 4.4 Protein autocorrelation function at 300 K, long time behavior 29

Figure 4.5 Protein-water cross-correlation functions 30

Figure 4.6 τ_{avr} of Φ_{PP} fits 32

Figure 4.7 τ_{avr} of Φ_{WW} fits 32

Figure 4.8 Frequency dependent complex dielectric constants. Protein dielectric constants at Ia) 170 K, Ib) 190 K, Ic) 210 K, Id) 250 K and Ie) 300 K; Water dielectric constants at IIa) 170 K, IIb) 190 K, IIc) 210 K, IId) 250 K and IIe) 300 K; Dielectric constants of the total solution at IIIa) 170 K, IIIb) 190 K, IIIc) 210 K, IIId) 250 K and IIIe) 300 K 35

Figure 4.9 Cole-Cole plots for a) protein, b) water, c) total system 36

Figure 4.10 Residue 75 of lysozyme. 37

Figure 4.11 Torsional angle trajectories ψ and χ_2 for residue 75 38

Figure 4.12 Water radial distribution functions around residue 75 at a) 170 K, b) 190 K, c) 210 K, d) 250 K and e) 300 K 41

Figure 4.13 Bulk water radial distribution functions at a) 170 K, b) 190 K, c) 210 K, d) 250 K and e) 300 K 42

LIST OF TABLES

Table 1.1 The static pair susceptibilities χ_{ij} and component susceptibilities χ_i of HIV1 zinc finger peptide in aquous solution[a] [13] 6

Table 1.2 The static pair susceptibilities χ_{ij} and component susceptibilities χ_i of an aquous Ubiquitin solution[a] [13] 7

Table 4.1 Fit Parameters for Φ_{PP}. A_1 + A_2 normalization is rescaled to 1 for easier comparison. 31

Table 4.2 Fit Parameters for Φ_{WW}. A_1 + A_2 normalization is rescaled to 1 for easier comparison.. 31

Table 4.3 Protein static dielectric constants 33

Table 4.4 Water static dielectric constants 33

LIST OF SYMBOLS AND ABBREVIATIONS

A_i .....Dimensionless multiplicative coefficient

a..... Acceleration

χ .....Susceptibility

D.....Electric displacement field

ε.....Permittivity

ε_0..... Permittivity of vacuum

ε_r.....Dielectric constant

$\bar{\varepsilon}$ ..... Static dielectric constant

E .....Energy

E..... Applied electric field

F .....Force on a rigid body object

J..... Current

M.....Dipole moment

m..... Mass of a rigid body

P ..... Polarization

q.....Charge

r ..... Generalized coordinate

T ..... Temperature

t .....Time

τ.....Relaxation time

U..... Potential energy function

V.....Volume

v.....Velocity

ω.....Frequency

x.....Distance

1. INTRODUCTION

Dielectric properties of proteins are of equal interest to theoreticians and experimentalists. Dielectric constant and conductivity of protein solutions and their dependence on conductivity are crucial on themselves, which renders calculation of these properties from computer simulations necessary. Static dielectric constant is also important due to its role in the Poisson–Boltzmann equation [1]. A static dielectric constant for protein and the dielectric medium around the protein is required to solve this equation, which itself is needed in the calculation of the electric field generated by the protein. Therefore, calculation of static and frequency dependent dielectric constants are both a hot research topic and a necessity for predicting other properties of the proteins. Moreover, calculation of dielectric properties presents a whole set of tools for analysis. Similarities and dissimilarities of these properties at different regions and dielectric correlations between the medium and the protein provide grounds to draw conclusions about the protein, the solvent around it and their interaction.

The dielectric reaction of a liquid to a frequency dependent external electrical field is well known [2, 3]. This theory, built on polarization and reorientation of individual molecules according to the external field, does not neatly apply to proteins [4]. Proteins include strongly polar and charged regions, which suggests that their reaction to an external field would be considerable. Yet, due to the long backbone and firm secondary and tertiary structures, reorientation of dipolar groups are limited and coupled. Thus, there are numerous theories regarding the dielectric response of proteins, and a large variety of static dielectric constant values have been reported [5-14]. Experimental verification of these values also present difficulties, since it is very hard to separate the response of the protein from the solvent around it. The dielectric properties of the protein itself and the whole solution containing it are two different parameters, and the former is not directly measurable. Furthermore, the counterions in the solution affect the outcome. Such difficulties have caused several different theories for

estimation of dielectric properties to be born. These theories are explained in Chapter 2. Following is a short summary of important work done on the subject, involving the aforementioned theories. Reader is advised to examine the corresponding sections of Chapter 2 for each publication, since the details of the theories used in each paper are given there.

In 1988, MD was not feasible yet; local static dielectric constants of BPTI were calculated from normal mode analysis in vacuo [5]. Inside the protein, local dielectric constants ranging from 1 to 20 were calculated from electronic polarization of atoms and orientational polarization of local dipoles.

Later, when short MD simulations became applicable, 50 ps long MD simulations were performed on trypsin in water [6]. The simulation used surface constrained all atom solvent model (SCAAS), the solvent around trypsin was divided into layers. According to the distance to the protein, the layers had unrestricted water molecules, then increasingly restricted water molecules and in the end an electrostatic continuum. The electrostatic interactions were not cut off. From these simulations, local static dielectric constant of different sites and static dielectric constant of water were calculated using two different approaches: *a)* Kirkwood-Fröhlich theory (see section 2.3.1) and *b)* using average electric field and polarization calculated from the simulation trajectory (see section 2.3.2). The calculated local static dielectric constants ranged from 3 to 20. Direct use of averaged field and polarization could not provide constants higher than 10. Kirkwood-Fröhlich theory led to constants above that value. The paper also includes dipole-autocorrelation functions for some of the sites. They have different characteristics, but share a common property: Decay to zero happens very fast, in about 14 ps.

Another paper was published the same year by another group [7]. This paper, too, investigates the local static dielectric constants, in the grounds that the biological function of a protein is highly dependent on the local variations in the dielectric properties. MD simulations of deca-alanine and cytochrome c were performed. Deca-alanine simuation was 150 ps after equilibrium, cytochrome c was 90 ps after equilibrium. The proteins were in vacuo. The force field was CharmM 19. The calculations were based on Kirkwood-Fröhlich theory (see section 2.3.1). Static

dielectric constants were calculated to be 3.3 for deca-alanine and 3.5 for cytochrome c. The susceptibilities of different residues were investigated and found to be varying by a factor of 4.

Two years later, MD simulations of length 1.4 ns for BPTI and 1 ns for lysozyme were carried out [8]. The proteins were in a solvent consisting of SPC/E or SPC model water. The force field was GROMOS. Electrostatic interactions were handled by a twin range method based on the Coulomb potential. Static and frequency-dependent dielectric constants were calculated for each protein. They applied the Neumann version of Kirkwood-Fröhlich theory (see section 2.3.1), which is based on one component (protein only), of the three component system (protein, water, ions). The main assumptions were that the protein is spherical and that the cross-correlation between protein and water is negligible. Prior to the calculations, the overall rotation (tumbling) of the protein was removed by a quaternionic fit, since 1 ns is not long enough to sample this rotation (see section 2.3.4). This removal of rotation and translation was done after the simulation was completed. Since the proteins were charged, the dipole moments depended on origin, the center of mass of the protein was chosen as the origin (see section 2.3.6.1). The static dielectric constants were calculated as 36 for BPTI and 10 for lysozyme. When the same calculations were carried out by leaving side-chains out of the dipole fluctuation considerations, the constants were found to be between 2 and 3. This is an expected outcome, since a protein has a low-dielectric core and a high-dielectric surface (which interacts with water). Autocorrelation functions of the protein dipole moments were fitted (not very well) to single exponential functions. From these fits, relaxation times of 1.8 ns for BPTI and 3.4 ns for lysozyme were calculated. Frequency-dependent dielectric constants were also calculated from these fits, found to have decayed to zero at around 10 Mhz. Cole-Cole plots were drawn.

In 1994, pK_a values of ionizable groups in proteins were calculated using the solution of Poisson-Boltzmann equation [9]. Static dielectric constants are a required input for this method. Although the convention is to input a low dielectric constant between 2 and 4 for the protein, this work reports that the best agreement with experiments is achieved when a much higher static dielectric constant, 20, is used. They

have concluded that this high constant is needed to incorporate conformational relaxation, which is not modeled elsewhere in this method.

Taking a larger step from their previous work [7], Simonson and Perahia performed a 1ns long MD simulation of ferro- and ferricytochrome c in 1995 [10]. Each protein was in a spherical volume of water molecules. Kirkwood-Fröhlich theory was applied in the calculations (see section 2.3.1). The side chains were found to have a large effect on dipole fluctuations of the whole protein, due to their fast motions. Being at the surface of the protein, they interact strongly with the surrounding water. Including the side chains, the static dielectric constants were calculated to be varying between 16 and 37. If the side chains are considered to be part of the solvent, the remaining core of the proteins appear to have the static dielectric constants of 4.7 for ferro- and 3.7 for ferricytochrome c. These findings are somewhat in accordance with the findings of Smith et al. in 1993 about the side chains and the core [8]. Commenting on these results, Simonson and Perahia argued that considering these side chains as part of the protein would be wrong, since this prevents the treatment of the protein as a homogenous dielectric material. The importance of local dielectric properties of the proteins was stated in the previous work. This paper also investigates variations in local static dielectric constants. The static dielectric constant in the inner half of the protein was found to be between 1.5 and 2. The suggestion about the Poisson-Boltzmann equation [1] in this work is to use these low dielectric constants for the proteins (the cores) and to consider the side chains as part of the solvent. This is in contradiction to the suggestion by Antosiewicz et al. in 1994 [9].

An MD simulation of the triple helical DNA strand $d(CG.G)_7$ was performed in the same year [15]. The length of the simulation was 1.115 ns. The ionic solution included 837 water molecules, 37 sodium ions and 16 chloride ions. The SPC/E water model was used. The force field was CharmM22. Electrostatic interactions were handled by the Ewald summation method. The system was conceived as a five component system: base / sugar / phosphate / water / ions, these components were treated just as in three component cases (protein / water / ions). Phosphate and ions components were charged, therefore their dipole moments were dependent on the origin. Center of mass of the DNA was chosen as the origin (see section 2.3.6.1). This also eliminates the contribution of DNA to the conductivity of the system. The static

dielectric constant was calculated from dipole moment fluctuations for each of the components. Cross-terms were found to be small, therefore they were neglected at the calculation of static dielectric constants. Only the component's own dipole moment fluctuations were considered for each part. The calculated static dielectric constants were 41.3 for water, 3.4 for bases, 2.0 for sugars, 33.0 for phosphate groups. The total static dielectric constant for the whole DNA was found to be 15.5. Then the fifth component, the ions, was treated in the same way as the others, and also a static dielectric constant was calculated for the ions instead of conductivity. The dipole relaxation time for SPC/E water was calculated as 9.7 ps.

After two more years, even longer simulations were feasible, and an MD simulation that lasted for 13.1 ns after equilibration was carried out on zinc finger peptide, a small (18 residues), neutral protein [12]. The system consisted of the protein, one zinc ion, two chloride ions and 2872 water molecules. The simulation was performed under periodic boundary conditions (a box-shaped simulation system). SPC/E water model was chosen and the united-atom CharmM19 force-field was used for non-water-water interactions. Electrostatic interactions were calculated using the Ewald summation technique. The authors claimed that use of Kirkwood-Fröhlich theory was not acceptable, and also the simulations can only lead to correct results under certain conditions. These conditions were laid out as rules (see section 2.3). They used a combination of linear response theory, phenomenological equations of matter and a computer-adapted dielectric theory to calculate static and frequency-dependent dielectric constants of protein and water and the conductivity of the ions. They have found that the contribution from the cross-correlation between dipole moments of protein and water components has an important contribution to the dielectric constant of the protein. The static dielectric constant of the peptide was calculated to be 15. The contribution from the cross-correlation term was 3 (if protein-water cross-term was neglected, the constant would have been found as 12). The static dielectric constant of the water was found to be 45, while pure SPC/E water has the constant 71, this difference in values was connected to the reduced mobility of the water molecules at the protein surface. The relaxation times of the protein were higher that the ones reported by Smith et al. [8], the main difference is that tumbling of the protein is included in this study, which is a dominant slow dielectric relaxation mode. Also different from Smith et al. [8], this work uses biexponentional fits, which provide two distinct relaxation

times and fit considerably better to the correlation functions. The frequency-dependent dielectric constants were shown to vanish at around 10^{-3}-10^{-2} ps^{-1}. The authors point to one possible problem with the method involving decomposition of linear response theory: each component must behave as a dielectric matter. As granularity decreases, the results become less related to those of a macroscopic dielectric.

In an extension to the work above, a new formalism called dielectric field equation (see section 2.3.3) is introduced, which allows combination of results from quantum mechanical, molecular dynamics and continuum electrostatics calculations that are executed on different parts of the system [13]. In addition to this, the zinc finger peptide trajectory from the earlier simulation [12] was re-analyzed, this time dividing the water molecules into three parts: first and second solvation shells, and bulk water. The separation was realized using Voronoi polyhedra. These were treated as different components and their behavior and contributions to the dielectric constant of the protein were investigated. The first shell was found to behave very differently indeed, but it was seen that the considerable contribution from the water component [12] was not mainly from the solvation shells, but bulk water had a serious contribution. This means that the coupling between the bulk water component (which consists of water molecules that are not immediately near the protein) and the protein is non-negligible. Table 1.1 shows the self- and cross-component susceptibilities, multiplied by 4π, so that each number corresponds to dielectric constant minus 1.

Table 1.1 The static pair susceptibilities χ_{ij} and component susceptibilities χ_i of HIV1 zinc finger peptide in aquous solution[a] [13]

χ_{ij}	P [b]	$S1$[c]	$S2$[d]	B [e]	χ_i [f]
P	10.6	-0.3	0.3	2.8	13.4
S1	-0.3	2.3	0.3	0.5	2.8
S2	0.3	0.3	3.9	1.4	5.9
B	2.8	0.5	1.4	32.5	37.2

[a] All susceptibilities are multiplied by 4π to facilitate comparison to component DCs.

[b] Protein, [c] First shell, [d] Second shell, [e] Bulk Water

[f] Obtained as the row sum of χ_{ij}, [g] Sum of all susceptibilities

A detailed analysis of the behavior of the three water partitions is presented, but the conclusion that stands out is that even though S1 and S2 behave differently, the main contribution of the protein-water cross-term comes from the bulk water and it heavily dominates the effect of S1 and S2. This shows that such a division of the water can provide a better understanding of the dielectric behaviour of the system, but does not offer a considerable improvement on the calculation of the dielectric constant of the protein.

The same group performed a 5 ns (after equlibration) MD simulation of the small protein ubiquitin (76 amino acids) in a cubic box with periodic boundary conditions the next year, 2000 [14]. SANDER module from AMBER 4.1 suite of programs was used, the force field was the Cornell et al. all atom force field [16]. SHAKE algorithm was used for bond lengths, and Particle Mesh Ewald was the method for electrostatic interactions. As in a former study, water was divided into three parts (S1, S2 and bulk) using Voronoi polyhedra [13]. Same investigations as in [13] was carried out on this simulation, and similar findings regarding the protein-water cross term was found.

Table 1.2 The static pair susceptibilities χ_{ij} and component susceptibilities χ_i of an aquous Ubiquitin solution[a] [13]

χ_{ij}	P [b]	$S1$[c]	$S2$[d]	B [e]	χ_i [f]
P	29.4 ± 1.1	-2.8 ± 0.4	1.0 ± 0.3	11.4 ± 1.2	39.0
S1	-2.8 ± 0.4	3.9 ± 0.1	1.4 ± 0.1	0.8 ± 0.4	3.3
S2	1.0 ± 0.3	1.4 ± 0.1	5.5 ± 0.1	5.1 ± 0.5	13.0
B	11.4 ± 1.2	0.8 ± 0.4	5.1 ± 0.5	65.8 ± 2.1	83.1

a All susceptibilities are multiplied by 4π to facilitate comparison to component DCs.

[b] Protein, [c] First shell, [d] Second shell, [e] Bulk Water

[f] Obtained as the row sum of χ_{ij}

In this paper, the following explanation to this phenomenon of anticorrelation in P-S1 and a higher than expected number as the P-B susceptibility term was presented by the authors: In this work and many of the other works in this field, a single solute with a

large dipole moment (protein or DNA, etc) interacts with the water molecules. To avoid artifical directing influences of this single solute, a simulation trajectory should include all possible orientations of the large solute. Therefore 5 ns is not enough as a simulation length and causes artifacts. This orientational diversity is realized by the presence of multiple, orientationally non-equivalent solutes. This explanation is a repetition of the last element of the list of rules the same group has established at 1997 (see section 2.3) [12]. In this paper these authors claim that the insufficient simulation length and the lack of sampling all of the rotations causes the protein-bulk cross-term to be this high. They support their reasoning by showing the distance-dependent Kirkwood g-factor for two different orientations of the protein, and concluding that the P-W crossterm depends strongly on the orientation. Neither the protein term nor the water term depend near as much on the orientation. The authors claim that with such short simulations, relaxation times (found from a bi-exponential fit) can be approximated, but the contribution of the protein-water cross-term contribution to the solution (and protein) dielectric constant cannot be determined.

This study focuses on the changes in the dielectric properties with temperature; therefore it is also necessary to present a background on the transition of the protein in the studied temperature region. The glassy relaxation phenomenon is the subject of active research. It has been shown that proteins experience a dynamical transition in the range ~190 – 220 K [17]. The protein is functional above the temperature of this transition. The temperature dependence of mechanical fluctuations has been investigated both experimentally by measuring average fluctuations of hydrogens under neutron scattering [17-21] and theoretically [22-27]. The transition is not observed in absence of water [28, 29] and the system experiencing the transition has been identified as both the protein and the solvent shell around it. Protein systems with hydrophilic solvents other than water, like glycerol, have been shown to experience transition as well [30]. Since the dynamical transition is dependent on the existence of a solvent, the protein – solvent interaction before, during and after the transition has been a point of focus [21, 27, 31, 32]. These works, together with the outcomes of previous studies on the systems without solvent, are proposing that the transition is triggered by the solvent. Until now, the dynamical transition of proteins had been only investigated using mechanical properties. Therefore the findings of these studies on the protein-solvent interaction were limited to the immediate surface of the protein. By using dielectric

properties, which are related to electrostatic forces with much longer ranges than mechanical interactions, as a tool to investigate, it is possible to better understand the phenomenon of dynamical transition in terms of the interaction between the protein and the solvent around it. This approach permits investigation of not only the interaction of the surface solvent molecules and the protein, but also the effect of this interaction on the whole solvent.

The scope of this work is calculation of dielectric properties of an aqueous lysozyme solution from an MD simulation, as in the summarized works in literature. Static and frequency dependent dielectric constants of the solution and its components will be obtained. As a novel approach, these properties will be calculated over a range of temperatures, which gives the opportunity to see how these parameters change with temperature and observe and analyze the dynamical transition using dielectric properties. Dependence of the dielectric parameters on temperature will also be used as a tool to analyze protein-water interaction and the dynamical transition of protein, which was shown to be around 195 K by mechanical analysis [22, 33-35]. The electrical analysis will be supported by further mechanical analysis such as torsion angles and radial distribution functions.

2. THEORETICAL BACKGROUND

2.1 Molecular Dynamics

Molecular Dynamics is a computational technique to simulate motions of many-body systems by integration of their equations of motion [36]. The trajectory of a system of particles allows calculation of its structural and dynamical properties. These motions in space and time are calculated using Newton's second law:

$$F_i = m_i \frac{d^2 r_i}{dr^2} \tag{2.1}$$

where F_i is the force acting on the atom, m_i is the atom mass, and r_i is the position vector of the atom. The force is the gradient of the potential energy U:

$$F_i = -\nabla_r U \tag{2.2}$$

U is a function of the positions of all atoms, and accounts for the sum of all interactions. This potential is calculated from a forcefield [37]. The forcefield used by NAMD, and therefore in the reported simulations, is the CHARMM forcefield [38]. It includes 2-, 3-, 4-body interactions, electrostatic interactions, and van der Waals interactions [39].

Trajectories of all particles in the system can be calculated from the derivative of the forcefield, according to Newton's second law of physics, which allows calculation of acceleration from the force. By numerical integration, velocities can be acquired from accelerations, and displacements can be acquired from velocities. A numerical approximation is necessary here, since there is no analytical solution to the equations of motion due to the complexity of the potential. This numerical integration is based on approximating the dynamical parameters by Taylor series expansions.

The numerical method of integration used in the reported simulations is the Verlet algorithm [40]. Acceleration is known from Newton's second law. One integration is needed for velocity and another one for position. Verlet algorithm updates the position, and then uses old and new positions to update the velocity. Position is written shifted forwards and backwards in time for equal amounts, *h*, and used together to solve for $x(t+h)$ and $x(t-h)$.

$$x(t+\Delta t) = x(t) + v(t)\cdot\Delta t + \frac{1}{2}a(t)\Delta t^2 + \frac{1}{6}\frac{d^3x(t)}{dt^3}\Delta t^3 + O(h^4) \tag{2.3}$$

$$x(t-\Delta t) = x(t) - v(t)\cdot\Delta t + \frac{1}{2}a(t)\Delta t^2 - \frac{1}{6}\frac{d^3x(t)}{dt^3}\Delta t^3 + O(h^4) \tag{2.4}$$

which leads to

$$x(t+\Delta t) = 2x(t) - x(t-\Delta t) + \frac{1}{2}a(t)\Delta t^2 + O(h^4) \tag{2.5}$$

2.2 General Theory of Dielectrics, Susceptibility and Permittivity

When an electric field is applied to a dielectric medium, current flows in this medium. The current can be separated into two parts: a conduction part, which accords to an actual current; and a displacement current part, which can be perceived as the elastic response of the medium to the applied field. Figure 2.1 explains this elastic response over an example relevant to the object of this study.

Polar molecules in a dielectric medium are oriented randomly without an applied electric field. When an external field is applied, the material is polarized; the dipole moments of the polar molecules will be oriented towards the applied electric field. This polarization creates an electric field opposing the applied field, therefore decreasing the effective electric field and increasing the capacitance of the parallel plates in the example of Figure 2.1.

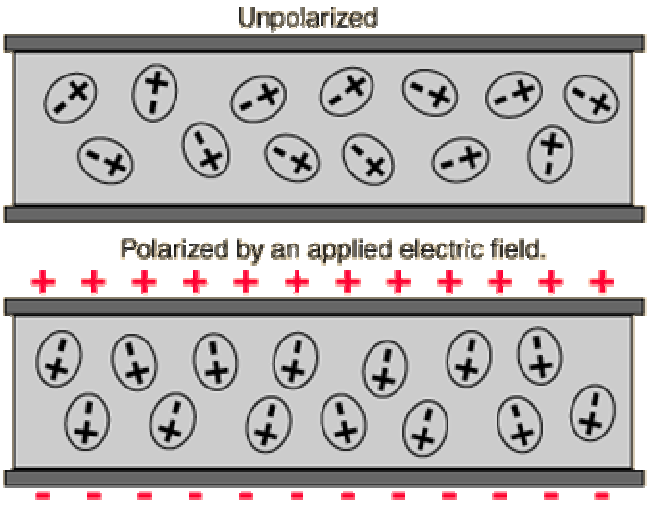

Figure 2.1 Response of a dielectric medium containing polar molecules to an applied electric field between parallel plates [41]

Here, two important terms, susceptibility and permittivity come up. Permittivity is the general quantity that describes how an electric field affects a dielectric medium and how that medium is affected by the electric field. It is a measurement of how easily the medium can polarize and reduce the effective electric field, when an external electric field is applied; therefore it is a measurement of how much of the external field is permitted through the dielectric medium. The electric susceptibility is directly related to the permittivity and is defined as the ability of a medium to be polarized by an external electric field. The very close definitions of the two parameters can be better explained by equations.

As mentioned above, the effect of an applied electric field **E** can be in two ways: Charge migration and dipole reorientation. Both these effects on the electrical charge distribution of the medium are accounted by **D**, electric displacement field. Permittivity is defined as the constant of proportionality relating the electric field to the electric displacement field:

$$\mathbf{D} = \varepsilon \mathbf{E} \tag{2.6}$$

where ε is permittivity. It is a scalar if the medium is isotropic and a 3x3 matrix if this is not the case.

Electric susceptibility relates the electric field to the dielectric polarization density **P** of the medium.

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E} \tag{2.7}$$

Here, ε_0 is the permittivity of free space (vacuum) and χ_e is electric susceptibility. It should also be noted that polarization is directly related to the dipole moment **M** as:

$$\mathbf{P} = \frac{\mathbf{M}}{V} = \frac{1}{V}\sum_{i=1}^{N} q_i \mathbf{r_i} \tag{2.8}$$

In this equation, *V* is the volume.

The permittivity of a medium relative to the permittivity of free space is called relative permittivity, or dielectric constant, and is denoted by ε_r.

$$\varepsilon = \varepsilon_r \varepsilon_0 \tag{2.9}$$

The susceptibility is related to the relative permittivity by

$$\chi_e = \varepsilon_r - 1 \tag{2.10}$$

From this equation follows that the electric susceptibility of vacuum is zero. The dielectric displacement **D** is related to the polarization density **P** as:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 (1 + \chi_e)\mathbf{E} = \varepsilon\, \mathbf{E} \tag{2.11}$$

Unlike vacuum, polarization of a dielectric medium depends on the frequency of the applied field, since the material cannot polarize instantly.

$$\mathbf{P}(\omega) = \varepsilon_0 \chi_e(\omega)\mathbf{E}(\omega) \tag{2.12}$$

with ω being the frequency of the electric field. This frequency dependence of the response is due to causality, which also makes permittivity a frequency dependent complex function.

$$\mathbf{D_0} e^{i\omega t} = \varepsilon(\omega)\mathbf{E_0} e^{i\omega t} \tag{2.13}$$

where $\mathbf{D_0}$ and $\mathbf{E_0}$ are the amplitudes of dielectric displacement and applied electric field. The static dielectric constant $\bar{\varepsilon}$ is defined as

$$\bar{\varepsilon} = \lim_{\omega \to 0} \varepsilon(\omega) \tag{2.14}$$

The high frequency limit is often denoted as ε_∞. The real and imaginary parts of the complex permittivity can be written as

$$\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega) \tag{2.15}$$

ε' is the real part and ε'' is the imaginary part related to the rate at which energy is absorbed by the medium.

2.3 Dielectric Theories on Solutions of Macromolecules

2.3.1 Kirkwood-Fröhlich Teory

This theory, originally developed by Kirkwood and Fröhlich [2, 42], and improved by others in time [43-46], calculates the dielectric constant of a dielectric sphere, surrounded by a continuum of uniform dielectric constant. The assumption is included in this critical step:

$$\frac{(\bar{\varepsilon} - 1)(2\varepsilon_{RF} + 1)}{3(\bar{\varepsilon} + 2\varepsilon_{RF})} = \frac{4\pi \langle \mathbf{P} \cdot \mathbf{e} \rangle}{3E_0} = \frac{\langle \mathbf{M} \cdot \mathbf{e} \rangle}{R^3 E_0} \tag{2.16}$$

R is the radius of the sphere, ε_{RF} is the dielectric constant of the environment.

The following is valid if the medium is isotropic:

$$\frac{\partial\langle M\cdot e\rangle_{E_0}}{\partial E_0}=\frac{\beta}{3}\left(\left\langle M^2\right\rangle_{E_0}-\langle M\rangle_{E_0}^2\right) \tag{2.17}$$

Using equations 2.16 and 2.17 together results in:

$$\frac{(\bar{\varepsilon}-1)(2\bar{\varepsilon}+1)}{9\bar{\varepsilon}}=\frac{\beta}{3R^3}\left(\left\langle M^2\right\rangle_{E_0}-\langle M\rangle_{E_0}^2\right) \tag{2.18}$$

This method can be accepted for homogenous solvents, but for macromolecules, its assumptions do not hold. Therefore the suitibility of this theory to dielectric constant calculations is very questionable, despite the fact that most of the earlier works on dielectric constants used this theory. Also, some of these early works neglect the effect of the surrounding water and use the theory as if the protein sphere is in vacuum (ε_{RF} = 1), which contributes further to the unreliability of the results.

2.3.2 Direct Approach

The static dielectric constant can be calculated directly from the average electric field and average polarization that result from the simulation [6]. To avoid the sphere in the continuum approach of Kirkwood-Fröhlich theory, <***E***> of the investigated region is calculated directly by numerical averaging and put into

$$\bar{\varepsilon}=1+\frac{4\pi\langle\mathbf{P}\rangle}{\langle\mathbf{E}\rangle} \tag{2.19}$$

<**P**>, the polarization vector is also averaged directly:

$$\mathbf{P}=\frac{\mathbf{M}}{V}=\frac{1}{V}\sum_i q_i\mathbf{r_i} \tag{2.20}$$

This approach can also be used to look at specific local sites for their static dielectric constant, but the results from these calculations appear to cover a large range

and give very low values of dielectric constants, even lower than Kirkwood-Fröhlich theory, which shows that this is not a very reliable approach.

2.3.3 Dielectric Field Equation (DFE)

DFE is used to unify the results of quantum mechanical (QM) / molecular mechanics (MM) / continuum electrostatics (CE) calculations. It provides a means to integrate findings from these different methods to one meaningful conclusion [13]. In this approach, the core or a specific site of the protein can be examined by QM, all of it by MM and the surrounding water by CE, while still being capable of combining their outputs for an overall result.

$$\mathbf{E}(\mathbf{r}) = \int_V d\mathbf{r}' \left\{ -\nabla_\varphi (\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') + \mathbf{T}(\mathbf{r} - \mathbf{r}') \mathbf{P}(\mathbf{r}') \right\} \tag{2.21}$$

Equation 2.21 is the dielectric field equation. $\rho(\mathbf{r})$ is charge density, $\mathbf{P}(\mathbf{r})$ is dipole density, $\varphi(\mathbf{r})$ is defined by

$$\varphi(\mathbf{r}) = \frac{1}{r} S(\mathbf{r}) \tag{2.22}$$

where $S(\mathbf{r})$ is a screening function present in computer simulations modifying the electrostatic potential. This $\varphi(\mathbf{r})$ can be viewed as Coulomb interaction. $\mathbf{T}(\mathbf{r})$ is the dipole-dipole tensor, the double gradient of the respective interaction potential.

$$\mathbf{T}(\mathbf{r}) = \nabla\nabla_\varphi(\mathbf{r}) \tag{2.23}$$

For each region (QM, MM and CE), the charge density and the dipole density are calculated differently. But the DFE (Eq. 3 above) is valid everywhere. The total electric field is given by the sum of the electric field of each region, calculated by DFE using the related charge density and dipole density equation, in addition to the homogenous contribution from the boundary conditions, and the external field (if there is any).

For MM,

$$\rho_{MM}(\mathbf{r}) = \sum_j q_j \delta(\mathbf{r} - \mathbf{r}_j) \tag{2.24}$$

$$\boldsymbol{\rho}_{MM}(\mathbf{r}) = \sum_j \boldsymbol{\mu}_j \delta(\mathbf{r} - \mathbf{r}_j) \tag{2.25}$$

$$\boldsymbol{\mu}_j = \sum_A q_A \mathbf{r}_{jA} \tag{2.26}$$

2.3.4 Adjustment of Dielectric Boundary Conditions

The calibration of dielectric boundary conditions (different from geometric boundary conditions such as transparent boundary conditions, TBC) can be done according to references [13, 47]. The standard implementation of the Ewald sum takes λ_{EW} = 1 and therefore ε_{EW} = ∞ which corresponds to a conducting medium (infinite DC). This is called *tinfoil boundary condition*s or *conducting boundary conditions*. By changing the cutoff radius r_c and the parameter η, as shown in the references, one can set $\varepsilon_{EW}^{eff} = \varepsilon_0$, which means that the dielectric constant of the boundary region is the same as that of the simulated system (and not infinite). There are also cases that the boundary constant is set equal to the water dielectric constant, again closer to reality, but not as close as the described method. Although these settings provide a more physical calibration, tinfoil boundary conditions are acceptable as well.

2.3.5 Linear Response Theory

In general, linear response theory [12, 14, 44, 48-52] states that the expectation values $\langle \tilde{\mathbf{O}}(\omega) \rangle$ of the frequency-components of an observable **O** are directly proportional to the frequency-components $\langle \tilde{\mathbf{E}}_0(\omega) \rangle$ of the external field:

$$\left\langle \tilde{\mathbf{O}}(\omega) \right\rangle = \chi_{OP}(\omega)\tilde{\mathbf{E}}_{\mathbf{0}}(\omega) \tag{2.27}$$

Here, the susceptibility depends on the coupling of **O** to the entire polarization **P** of the system (as defined in section 2.2).

$$\chi_{OP}(\omega) = \frac{V}{3kT}\int_0^\infty \left(\mathbf{O}(0)\dot{\mathbf{P}}(-t)\right) e^{-i\omega t}\, dt \tag{2.28}$$

The observable **O** can be one of the three components that exist in a protein solution: Protein, water, or ions (if there are ions in the solution, which do exist to neutralize the total charge if the protein is not neutral). Since the dielectric constant of the protein is seeked, **O** will be $\mathbf{P_p}$ (the sub P is for protein, W for water, I for ions). Since total polarization **P** includes all the three components, the susceptibility of the observed component, which is directly related to its dielectric constant, is affected by the autocorrelation of the observable, and its correlations with the other two components.

$$P(t) = \frac{1}{V}\left[M_W(t) + M_p(t) + M_I(t)\right] \tag{2.29}$$

According to this, derivation of polarization of the protein follows as:

$$\left\langle \tilde{\mathbf{P}}_{\mathbf{P}}(\omega) \right\rangle = \chi_{P_P P}(\omega)\tilde{\mathbf{E}}_{\mathbf{O}}(\omega) \tag{2.30}$$

$$\begin{aligned}\chi_{P_P P}(\omega) = \frac{1}{3VkT}\Big[&\left(\mathbf{M_P}(0)\mathbf{M_W}(0)\right) - i\omega\int_0^\infty \left(\mathbf{M_P}(0)\mathbf{M_W}(t)\right)e^{-i\omega t}\, dt \\ &+ \left(\mathbf{M_P}(0)\mathbf{M_P}(0)\right) - i\omega\int_0^\infty \left(\mathbf{M_P}(0)\mathbf{M_P}(t)\right)\times e^{-i\omega t}\, dt \\ &+ \int_0^\infty \left(\mathbf{M_P}(0)\,\mathbf{J_I}(t)\right)e^{-i\omega t}\, dt\Big]\end{aligned} \tag{2.31}$$

Here, $\mathbf{J_I}$ is the current due to ions,

$$\dot{M}_I = J_I \tag{2.32}$$

The use of this theory in obtaining dielectric constants is as follows: The relation between the internal electric field and the external field is

$$\frac{\tilde{E}(\omega)}{\tilde{E}_0(\omega)} = \frac{2\varepsilon_{RF}+1}{2\varepsilon_{RF}+\varepsilon_W(\omega)+\varepsilon_P(\omega)-1} \tag{2.33}$$

Applying this to the linear response theory above, the following equation is obtained:

$$\frac{\varepsilon_P(\omega)-1}{4\pi} = f(\omega)\chi_{P_PP}(\omega) \tag{2.34}$$

and

$$\frac{\varepsilon w(\omega)-1}{4\pi} = f(\omega)\chi_{P_wP}(\omega) \tag{2.35}$$

where

$$f(\omega) = \left[1 - \frac{4\pi\left|\chi_{p_wp}(\omega)+\chi_{p_pp}(\omega)\right|}{2\varepsilon_{RF}+1}\right] \tag{2.36}$$

The dielectric constant equation for water and conductivity equation for ions are similar to this. $f(\omega) = 1$ if $\varepsilon_{RF} = \infty$ and this is true in an ideal implementation of Ewald sum (see section 2.3.4). In this case,

$$\varepsilon_P(\omega) = 4\pi\chi_{P_PP}(\omega)+1 \tag{2.37}$$

So, calculation of the susceptibility, which results from the calculation of correlation functions between the three components, directly leads to the dielectric constant.

2.3.6 Handling Charged Proteins

For any of these calculations, it must be noted that the whole system must be neutral. This poses no problems as Molecular Dynamics simulations also work under this restriction.

Also, the dipole moments of the systems are independent of the origin only if each of them are neutral. This creates a problem in the case of charged proteins. Then, the components P and I are not neutral. Ions have the total opposite charge of the protein to obtain neutrality of the whole system, but since their current $\mathbf{J_I}$, the derivative of $\mathbf{M_I}$, is in the calculations, they don't present any problem of origin. The charged protein however, does. There are two possible ways to act in this case.

In the center of mass method, the dipole moment of the charged protein is calculated by choosing its center of mass as the origin of the system [8, 11]. For each timestep, the dipole moment is calculated from the center of mass of the protein at that timestep.

Loèffler, Schreiber and Steinhauser have developed an alternative method [12]. This method involves reducing the net charge of the protein to zero by subtracting an equal amount of small charge from the partial charge of each atom. To compensate this subtracted charge, a pseudo-ion with a charge equal to the total subtracted charge (which is the net charge of the protein) is added as a new ion to the list of ions. To keep the total dipole moment of the system, the position of this ion has to be the geometric center of the protein. This way, both protein and ions components become neutral and the dipole moments are independent of the origin. Since the subtracted charge is very small for each atom, the calculations are not affected.

The charge of the protein contributes to the current, as the geometric center slowly moves. This is acceptable, since a protein with a net charge is a giant ion, but it's dielectric relaxation dominantly consists of oriental relaxation and not transformation, therefore its real relaxation is governed by the dielectric constant and not conductivity, so that this redistribution of charges between protein and ions is suitable to the calculations.

As an example, the protein in this study, lysozyme, has 1968 atoms and has a charge of +8, so 8/1968 is subtracted from the partial charge of each protein atom, and a pseudo ion with the charge +8 is created at the geometric center of the protein. This is done for each timestep.

The pseudo ion is included in the ions component and the protein component is the -now neutral- protein.

2.3.7 Simulation Rules for Dielectric Calculations

A set of rules for simulations with the purpose of dielectric relaxation calculations are determined by Loèffler, Schreiber and Steinhauser [12]. If these conditions are not satisfied, the reliability of the results are questionable. It is also imperative to state that these rules were set at 1997, and some of these may become obsolete in time (i.e. a theory better suited than linear response theory may take its place), but they keep their validity to this day.

(i) Boundary Conditions: Since dielectric properties are macroscopic, periodic boundary conditions are necesary: Box-shaped simulation systems (toroidal/periodic boundary conditions) or a hypersphere are acceptable only.

(ii) Treatment of Electrostatic Interactions: Neither Coulomb potential with cutoff (neither switched nor gradual), nor full Coulomb potential (without cutoff) are acceptable for treatment of electrostatic interactions. The following methods are acceptable: Ewald summation, other lattice summation methods, equivalent methods (such as Particle Mesh Ewald or Particle-Particle Particle Mash Ewald) and reaction field methods.

(iii) Theory: The combination of (i) linear response theory (see section 2.3.5), (ii) macroscopic definiton of dielectric properties (see section 2.2) and (iii) a computer adapted version of dielectric theory (see section 2.3.5) gives

the most successful results in three component systems. Kirkwood-Fröhlich theory is not reliable.

(iv) Simulation Length: All dielectric relaxation modes of a protein have to be sampled for truly meaningful results, including overall rotation. The simulation has to be long enough to sample tumbling. This length depends on the size of the protein, but at least 15 ns are required for even small proteins.

2.3.8 Calculation of Susceptibility Using Computer Adaptive Linear Response Theory

This section explains the details of calculating dielectric susceptibility (and dielectric constant, since the two are closely related by equation 2.10) based on the linear response theory. The susceptibility $\chi(\omega)$ of an object is obtained from the time correlation function $\Phi(t)$ of that object, according to equation 2.39. This equation is valid for tinfoil boundary conditions, which are used in the reported simulations.

$$\Phi(t) = \langle \mathbf{M}(t) \cdot \mathbf{M}(0) \rangle \tag{2.38}$$

$$\chi(\omega) = \frac{1}{3Vk_BT} \angle[-\dot{\Phi}(t)] \tag{2.39}$$

T is the temperature, k_B is the Boltzmann constant. $\angle[f]$ is the Fourier-Laplace transform function given as

$$\angle[f] = \int_0^\infty dr e^{-i\omega t} f(t) \tag{2.40}$$

The notation $\dot{\Phi}(t)$ corresponds to the first derivative of the time correlation function with respect to time.

As equations 2.15 and 2.40 suggest, susceptibility is a complex parameter.

$$\chi(\omega) = \chi'(\omega) + i\chi''(\omega) \tag{2.41}$$

It is possible to write the real and imaginary parts of the susceptibility from equation 2.40 as

$$\chi'(\omega) = \chi(\omega = 0) - \left(\frac{\omega}{3Vk_BT}\right)\mathrm{Im}\{\angle[\Phi]\} \tag{2.42}$$

$$\chi''(\omega) = \left(\frac{\omega}{3Vk_BT}\right)\mathrm{Re}\{\angle[\Phi]\} \tag{2.43}$$

For ω=0,

$$\chi_0 = \chi(\omega = 0) = \frac{\Phi(0)}{3Vk_BT} = \frac{\langle \mathbf{M}^2 \rangle}{3Vk_BT} \tag{2.44}$$

Total susceptibility is the sum of susceptibilities of all components. It should be noted that the ions component corresponds to conductivity instead of a dielectric constant. The way to calculate the susceptibility of each component is given in equation 2.31.

Due to the noise involved in correlation functions gathered from simulations, a curve fit has to be done on the acquired Φ functions to be able to apply equation 2.39. A biexponential fit and a stretch exponential fit are the best options considering wellness of fit and least loss of information. Biexponential fit has been shown to be sufficient for protein solutions and easier to calculate.

$$\frac{1}{3Vk_BT}\Phi(t) \approx \Phi_{fit} = A_1 e^{-\frac{t}{\tau_1}} + A_2 e^{-\frac{t}{\tau_2}} \tag{2.45}$$

Robustness of fitting was increased by adding an extra constraint:

$$A_2 = \chi_0 - A_1 \tag{2.46}$$

This normalizes the fitted correlation function to χ_0. Using this fit, the Fourier-Laplace transform can be completed to obtain

$$\chi'(\omega) = \chi(0) - \omega\left(\frac{A_1\tau_1^2\omega}{1+\tau_1^2\omega^2} + \frac{A_2\tau_2^2\omega}{1+\tau_2^2\omega^2}\right) \tag{2.47}$$

$$\chi''(\omega) = \omega\left(\frac{A_1\tau_1}{1+\tau_1^2\omega^2} + \frac{A_2\tau_2}{1+\tau_2^2\omega^2}\right) \tag{2.48}$$

The fit aims to remove the noise from the simulation to make it possible to analytically carry out the Fourier-Laplace transform. Since individual components are directly accessible, a simple ansatz such as biexponential decay is enough to convey all information. More complex functional forms do not present a higher level of illumination.

3. SIMULATION DETAILS

In this study, the dielectric properties of an aqueous Hen Egg White Lysozyme solution were examined. 2 ns long MD simulations were run at 150 K, 170 K, 190 K, 210 K, 230 K, 250 K, 270 K, 290 K and 300 K. The simulation at 300 K was prolonged to reach 4 ns. These simulations were prepared and realized using NAMD, Not Another Molecular Dynamics. 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 [39]. The initial structure of the protein was taken from the Protein Data Bank (PDB) with code 6lyz.pdb [53]. Hydrogen atoms are not included in the PDB file, because they are not resolved in X-Ray crystallography. These missing hydrogens were added and the protein was put in a box filled with 2769 TIP3 water molecules. TIP3 model was chosen for it is better suited to the forcefield used. In Visual Molecular Dynamics software (VMD), a 54 Å x 54 Å x 54 Å box was created. The thinnest layer of water was fixed at 5 Å. To neutralize the solution, eight chloride ions were added into the solvent. Periodic boundary conditions were set. Electrostatic interactions were handled by Particle-Mesh Ewald method (PWE). Tinfoil dielectric boundary conditions were used.

The system was energy minimized by 5000 conjugate gradient iterations. All bonds of protein and water molecules were constrained by RATTLE algorithm. Integration algorithm was velocity Verlet. The systems were equilibrated for 500 ps with a timestep of 2 fs. Temperature was kept constant by direct velocity scaling during this equilibration. For the data collection part of the simulation, temperature was controlled by a temperature coupling method. Data were recorded every 2 ps.

4. RESULTS AND DISCUSSION

In this work, protein-water interactions at different temperatures were studied via dielectric properties. The rules pointed out in section 2.3.8 were followed. Linear response theory (see sections 2.3.5 and 2.3.9) was adopted in calculations and Pseudo-ion method (see section 2.3.6) was implemented to handle the charges on the protein.

4.1 Dielectric Relaxation: Correlation Functions

The most important terms in the calculation of frequency dependent dielectric constants are the time correlation functions Φ. There are three components in the solution: Protein, water and ions. Total dipole moment autocorrelations of these components (Φ_{PP}, Φ_{WW} and Φ_{JJ}) and cross-correlations between them are all contributing to the dielectric constant of the solution. Dielectric constant of one component includes its autocorrelation and its cross-correlations with other components according to the equation 2.31. Among these, the highest contribution comes from the autocorrelation function.

4.1.1 Autocorrelations

Figures 4.1 and 4.2 show the normalized autocorrelation functions of protein and water components as temperature is increased. The first noticeable difference is that the relaxation of the total dipole moment of the protein component is much slower compared to the relaxation of water. This is clearly expected, as the protein, huge compared to water molecules, cannot reorient nearly as fast as water molecules. The autocorrelation functions of the ion component are even faster. For the ion component, the effective contribution comes as currents, as shown in equations 2.31 and 2.32. The

autocorrelation of currents are very fast compared to the other autocorrelation functions of the solution, and their cross-correlations have a negligible contribution to the calculation of dielectric constants. A current autocorrelation function (from the 290 K simulation) is given in Figure 4.3 to exemplify to their fast decay.

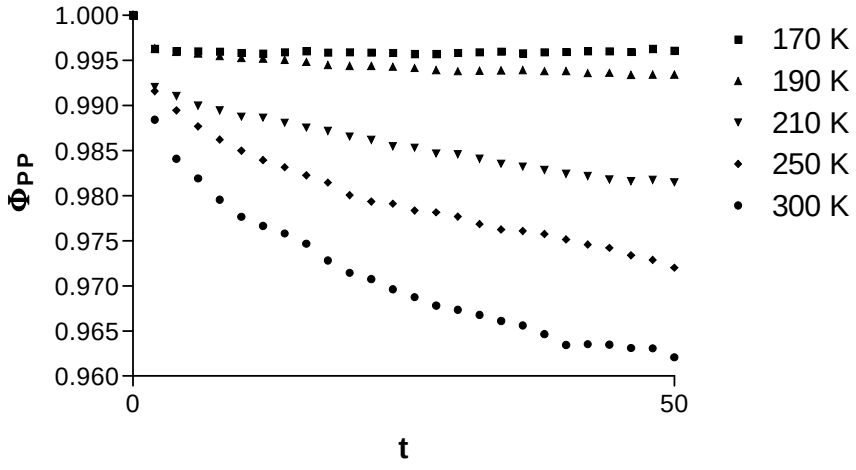

Figure 4.1 Protein component total dipole moment autocorrelation functions $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$

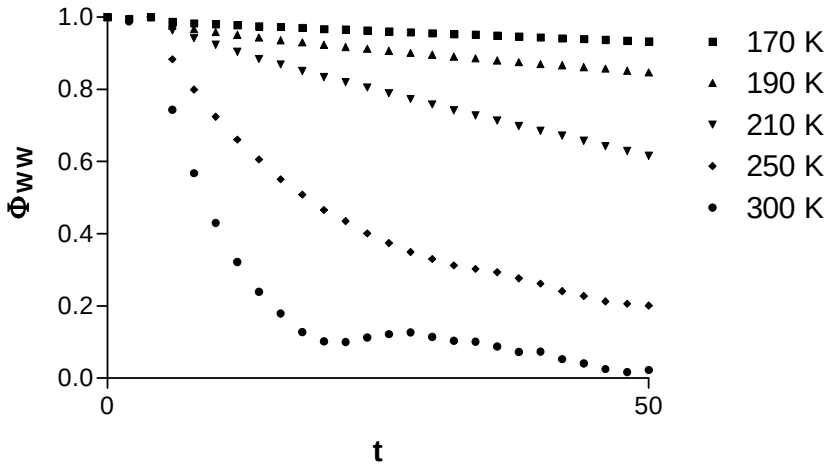

Figure 4.2 Water component total dipole moment autocorrelation functions $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$

These correlation functions are shown up to 50 ps, since this short time behavior samples the decay best. At longer time intervals, noise increases and discussion becomes less meaningful. A longer time behavior for protein autocorrelation obtained from the 8 ns MD runs at 300 K is provided in Figure 4.4.

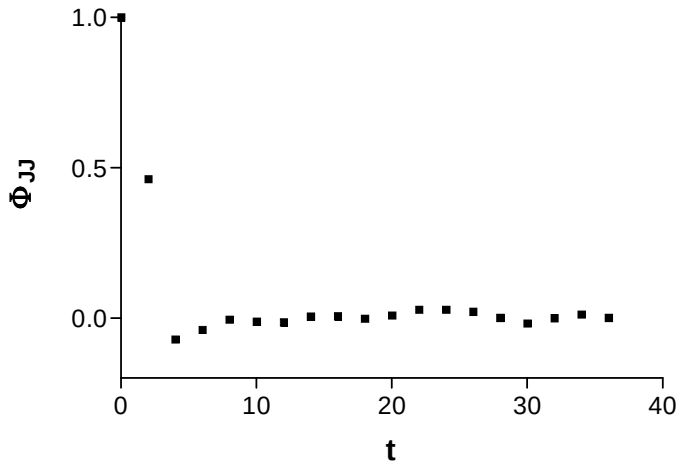

Figure 4.3 Current autocorrelation function at 290 K

Water autocorrelation shows increasingly faster decay as expected. With increasing temperature, mobility of the water molecules increase and the dipole moments of the water molecules can reorient much faster. The protein autocorrelation functions present interesting results. The 150 K, 170 K and 190 K group shows almost no relaxation at all. At higher temperatures than that, a trend of faster decay with increasing temperature is seen. It is essential to note that the mentioned trend is not as clear and neat as in the water component, since water dipole moment is the total moment of thousands of identical small molecules and therefore perturbations are averaged out, whereas the protein dipole moment belongs to one large macromolecule with different residues and partial charges at every contributing location. The observations point out to a transition around that temperature, but are not enough solely to prove it. Previous MD simulations on lysozyme showed a protein dynamical transition temperature of ca. 195 K from three different methodologies [34, 35]: (i) average fluctuations of C_α atoms in space, (ii) the stretch exponents of fits to the relaxation of displacement vectors of C_α atoms and (iii) heat capacity data [22, 33]. Note

that these findings are based on the mechanical properties of the system as opposed to the electrical properties studied in the current work.

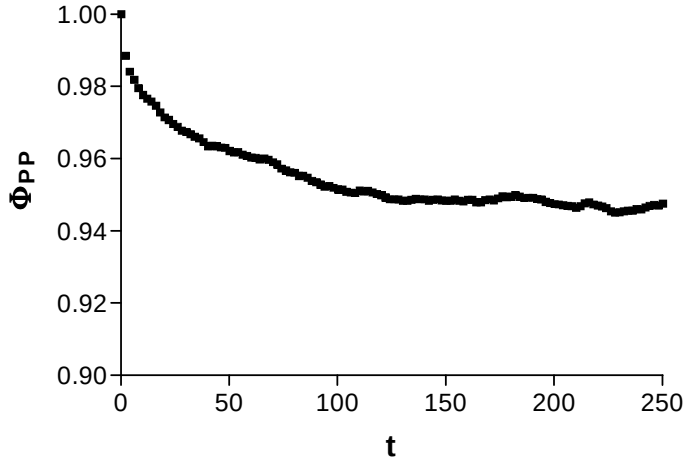

Figure 4.4 Protein autocorrelation function at 300 K, long time behavior

4.1.2 Cross-Correlations

Cross-correlations have very small contributions to the dielectric constants of each component [12-14]. Nevertheless, examination of these functions could give insight to the protein-water interaction. Cross-correlations of the ion component are completely insignificant and negligible. Though still having a very small contribution, the most significant cross-correlation function is protein-water cross-correlation, Φ_{PW}, which is shown for a number of temperatures in Figure 4.5. Only few temperatures are shown, since there are not significant differences between cross-correlations of different temperatures. Over all temperatures, the general trend of Φ_{PW} is fluctuation about 1. It is expected to experience a very slow decay, but the high levels of noise obstruct retrieval of this information. As stated before in literature [14], simulations of length on the order of the simulations reported in this work cannot sample the extremely slow tumbling mode of the protein, and therefore the contribution of Φ_{PW} to the dielectric constant values cannot be calculated correctly. However, even if it was possible, the

contribution is very small, at ca. 3. This value is most probably smaller than the noise present in the dielectric constant values over the temperatures.

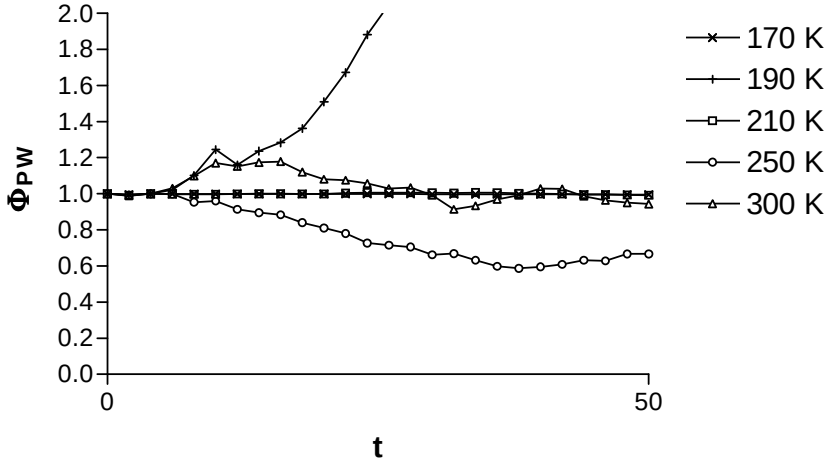

Figure 4.5 Protein-water cross-correlation functions

4.1.3 Double Exponential Fit

The time correlation functions are fit to a biexponential function with three independent variables (see section 2.3.9). The fit parameters for protein and water autocorrelation functions are given in Tables 4.1 and 4.2. In these tables, the first column shows how many timesteps were taken for the fit, which is different for some of the temperatures, since the low temperature cases proved harder to fit. This is expected, since the relaxation is extremely slow, even almost nonexistent at these temperatures. τ_{avr} is found by adding A_1 times τ_1 and A_2 times τ_2. It is the area below the correlation function graph – the average relaxation time. R^2 shows wellness of fit. These fits analyze the decay of these functions in two modes with different relaxation times, τ_1 and τ_2. τ_1 is the fast mode and is the dominant mode in water autocorrelation function as the ratio of A_1s to the according A_2s show. τ_2 is the slow mode and is dominant in protein. Figures 4.6 and 4.7 show change of the average relaxation times with temperature.

Table 4.1 Fit Parameters for Φ_{PP}. A_1 + A_2 normalization is rescaled to 1 for easier comparison. First column shows the number of timesteps used in fit.

Protein							
timesteps	T	A_1	τ_1	A_2	τ_2	τ_{avr}	R^2
10	170	0.0037	0.03	0.9963	22061	21979.4	0.9989
14	190	0.0037	0.75	0.9962	11120	11078.1	0.9989
14	210	0.0083	0.93	0.9917	3740	3708.9	0.9978
50	230	0.0131	1.96	0.9869	2564	2530.4	0.9965
50	250	0.0133	3.20	0.9867	3323	3278.9	0.9928
50	270	0.0178	2.15	0.9822	2640	2593.1	0.984
50	290	0.0162	3.09	0.9839	2039	2006.1	0.9967
50	300	0.0204	3.51	0.9796	2568	2515.7	0.9897

Table 4.2 Fit Parameters for Φ_{WW}. A_1 + A_2 normalization is rescaled to 1 for easier comparison. First column shows the number of timesteps used in fit.

Water							
timesteps	T	A_1	τ_1	A_2	τ_2	τ_{avr}	R^2
20	170	0.0113	0.98	0.9887	772.8	764.10	0.9994
50	190	0.0330	3.90	0.9671	346.2	334.92	0.9993
50	210	0.7571	66.90	0.2429	2508	659.84	0.9973
50	230	0.7709	27.89	0.2291	2565	609.14	0.9991
50	250	0.8260	15.52	0.1740	2568	459.65	0.9986
50	270	0.8730	10.33	0.1270	2568	335.15	0.9992
50	290	0.8461	6.35	0.1539	2576	401.82	0.9893
50	300	0.9430	6.31	0.0570	2566	152.21	0.9876

Following points are interesting to discuss. Protein average relaxation time makes a steep dive until 210 K, and starts fluctuating around 3000 ps above that. This is in accordance with the virtual lack of relaxation seen in the autocorrelation functions under 210 K. Decay in 150 K data is almost nonexistent so that it could not be fit to a biexponential decay. The others allow a fit, but only from a smaller set of timesteps. It is clear that the double exponential fit model does not describe the frozen states at the glassy region. The relaxation times are near values reported in literature for simulations of lysozyme [8]. At temperatures below 210 K, fits to water data are also hard and show a τ_2 dominance. Whereas these τ_2s are faster than protein's, they are still long compared to τ_1 from water at higher temperatures. This is due to hardness of fitting where the decay is very slow. The average relaxation times of water show a noisy, but more or less linear decay trend, with the most significant departure at 190 K. Perhaps the most interesting result is that the relaxation times of fast mode (τ_1) in protein at higher temperatures (around and above 250 K) are very close to the (dominant) fast mode

relaxation times of the water component. The slow mode of the protein must be including the overall rotation, or tumbling, of the protein. It is possible that the fast mode is mostly due to side chains, which heavily interact and move with water molecules at these temperatures, as will be shown in section 4.3. This proposed explanation, as given before in literature [12], fits to the matching τ_1 values as the relaxation times of water molecules and side chains could be close, and in is agreement with further results that will be presented. These interpretations of fit parameters only give a qualitative general picture. The real information these fits carry is exposed when the complex dielectric constants are calculated.

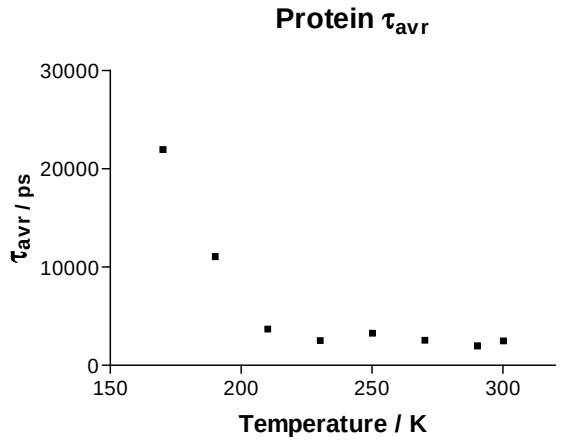

Figure 4.6 τ_{avr} of Φ_{PP} fits

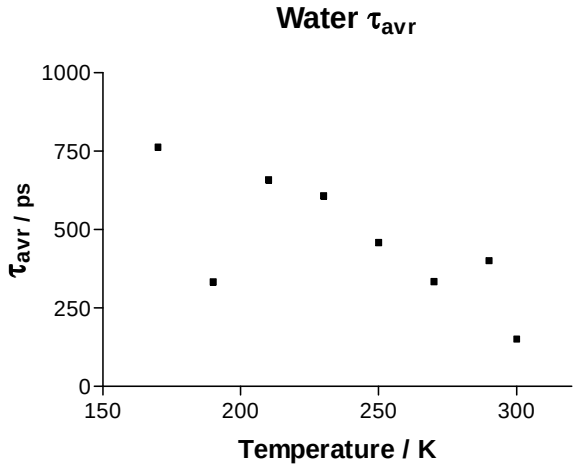

Figure 4.7 τ_{avr} of Φ_{WW} fits

4.2 Dielectric Constants

4.2.1 Static Dielectric Constants

The static dielectric constants of the protein from different temperatures is given in table 4.3. As expected, the static dielectric constant does not change much with temperature, except noise fluctuations. These noise fluctuations are bigger than the contributions of the crossterms, therefore the cross-terms were not included in the calculation of static and frequency dependent dielectric constants.

Table 4.4 lists static dielectric constants of water at different temperatures. The expected decay in accordance with theoretical and experimental data [54] is seen. This lowering of static DC is due to higher fluctuations in water molecules at higher temperatures, which reduces their effectiveness to stay polarized under a static electric field. The larger values for water in comparison to protein is also expected. The constants are smaller than that of TIP3 water, which is around 80. This is due to their slower relaxation, which is natural because of the large protein included in the solution.

Table 4.3 Protein static dielectric constants

T / K	DC
150	30.2
170	25.3
190	27.9
210	16.6
230	14.6
250	20.7
270	15.7
290	23.0
300	21.5

Table 4.4 Water static dielectric constants

T / K	DC
150	150.3
170	110.1
190	90.3
210	93.5
230	81.9
250	63.4
270	61.5
290	51.1
300	51.5

4.2.2 Frequency Dependent Complex Dielectric Constants

The frequency dependent dielectric constants of the protein, with their real and imaginary parts drawn separately, are shown in Figure 4.8. The shapes of the curves are not changed over the temperatures, but the decay of the real part (and the peak of the imaginary part) is shifted to higher frequencies, which is expected, since at higher temperatures, fluctuations are larger and the increased mobility allows to respond to higher frequencies. Figure 4.8 also shows the frequency dependent dielectric constants of water and total frequency dependent dielectric constants of the solution, both of which again show the slight shift in the peaks. Total dielectric constants carry both transitions of protein and water, which is normal, since they are the these two term were terms added to reach it. The terms from cross-correlations are insignificant and do not truly change the shape of the curves. The most important feature in these graphs is the nick on the water curves at the frequency of protein transition. The two step decay in the real part (or the two peaks in the imaginary part) in the total dielectric constant is obvious, since both protein and water terms are added in the process of calculating this constant. However, the dielectric constants of water are solely calculated from water molecules, yet at some temperatures they show a small nick in the real part and a small peak in the imaginary part at the frequency of protein transition. Moreover, this nick starts to appear only after 190 K, which suggests that there is a transition between 190 and 210 K, after which an interaction starts between the protein and water molecules. This can be seen better in the Cole-Cole plots drawn from these curves.

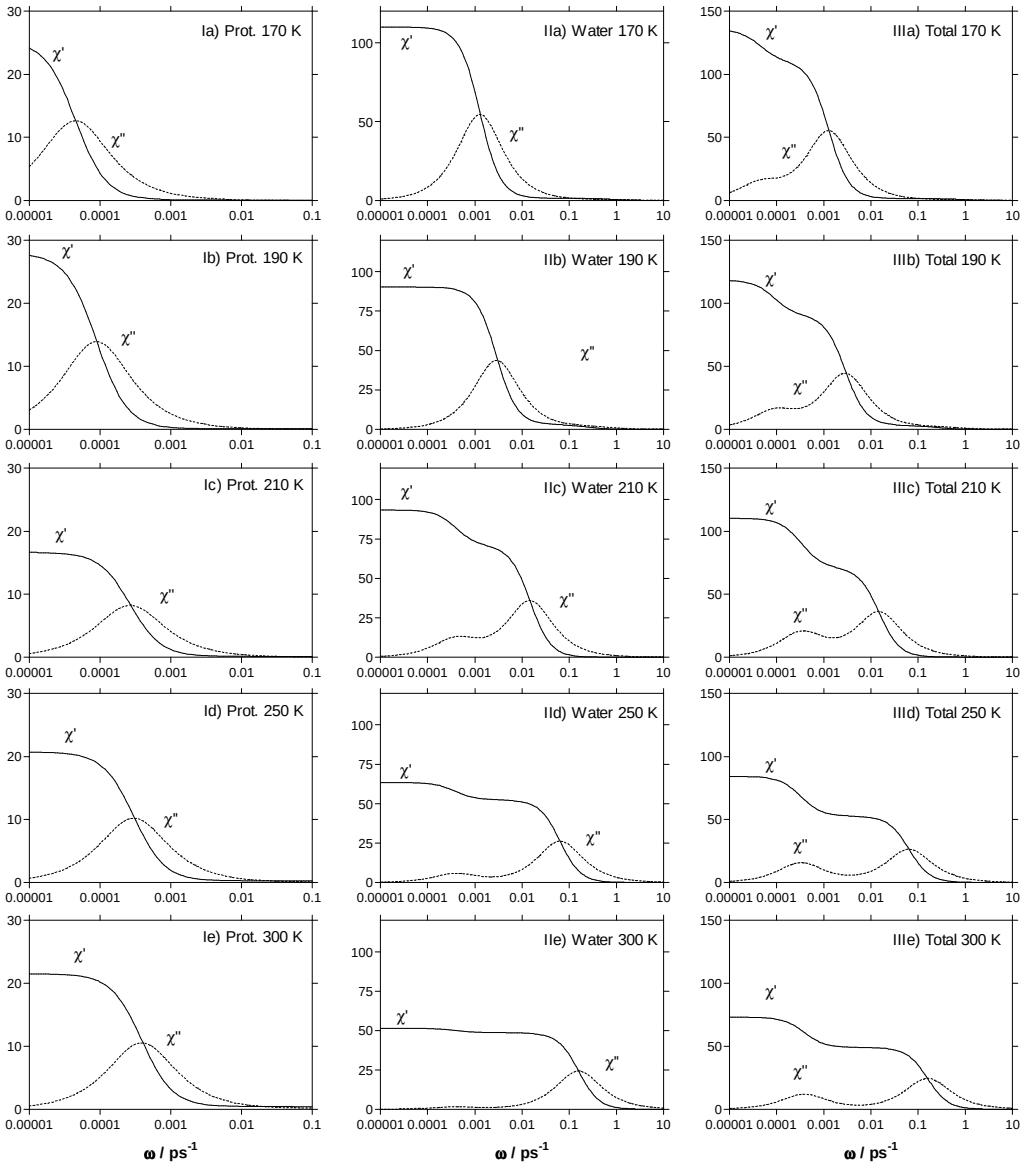

Figure 4.8 Frequency dependent complex dielectric constants. Protein dielectric constants at Ia) 170 K, Ib) 190 K, Ic) 210 K, Id) 250 K and Ie) 300 K; Water dielectric constants at IIa) 170 K, IIb) 190 K, IIc) 210 K, IId) 250 K and IIe) 300 K; Dielectric constants of the total solution at IIIa) 170 K, IIIb) 190 K, IIIc) 210 K, IIId) 250 K and IIIe) 300 K

4.2.3 Cole-Cole plots

The real and complex parts of the dielectric constant correspond to energy gained and energy lost by the system respectively. Cole-Cole plots are obtained by plotting the imaginary part of the complex dielectric constants against its real part. These curves present a better opportunity to see the transition mentioned in the previous sections. Figure 4.9 shows the Cole-Cole plots drawn for protein, water and the total system. The nicks in Figure 4.8 water dielectric constants can be seen clearer in Figure 4.9b. Starting at 210 K, the half circles of water component start to have an addition of a smaller lobe with the diameter of the half-circle of the protein at the same temperature. The presence of two merged lobes is trivial at the total solution Cole-Cole plot, since both protein and water components are included in it, but this appearance of protein lobes in water component definitely points out that the water component is affected by the protein, but only at and after 210 K. Moreover, the total system contains not three but two of these lobes. The lobes of the water behavior merge with that of the protein on such a way

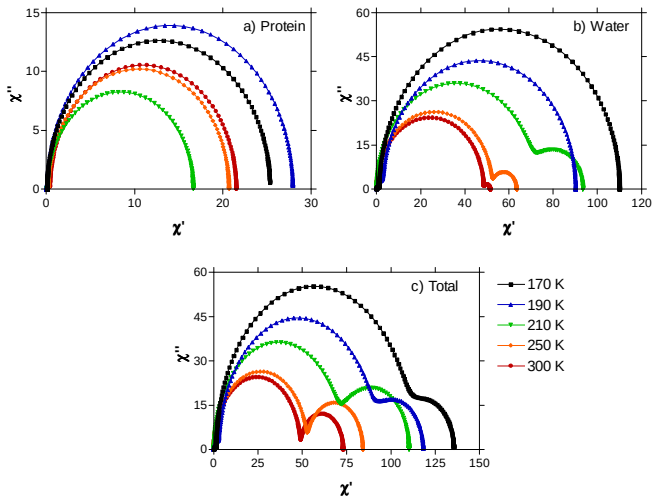

Figure 4.9 Cole-Cole plots for a) protein, b) water, c) total system

Figure 4.9b is clear evidence that there is a transition at around 200 K, after which an active protein-water interaction starts. This can be tied with the protein's lack of function below these temperatures. At these low temperatures, proteins show glassy behavior as discussed earlier. Figures 4.8 and 4.9b suggest that this transition from non-

functional to functional protein is connected with the onset of interaction with the water molecules surrounding the protein. This interaction is related to the side chains, which start to move by changing their torsion angles, as previously exemplified for BPTI [33]. To investigate closely on this, several residues with side chains at the surface of the protein were chosen and their interactions with water molecules around them were examined.

4.3 Side Chains and Water Molecules

4.3.1 Mobility of Side Chains

In an effort to show how the mobility of the side chains change with temperature, backbone and side chain torsion angles were calculated. Figure 4.10 shows residue 75, which exemplifies a surface side chain. Torsional angle trajectories at different temperatures for this residue are reported in Figure 4.11. The ψ angle is the torsion angle on the backbone, representing local fluctuations. At temperatures below unfolding process, there are no jumps in this angle, as the average structure of the protein is intact. The fluctuations around the mean, however, increase as temperature increases.

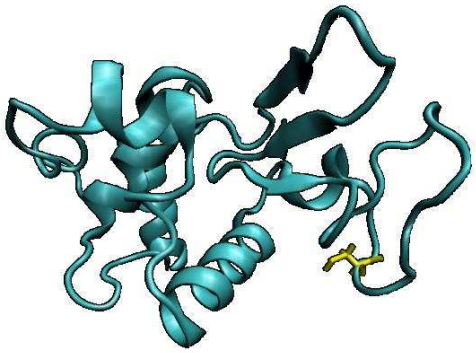

Figure 4.10 Residue 75 of lysozyme. The protein is shown in blue cartoon presentation, only the side chain of residue 75 is presented among all residues, colored yellow

χ_2 angles (torsion angle between C_α and C_β) are not fixed, since a rotametric jump on a χ_2 angle means a change in the orientation of the side chain, which is not restricted by the firm tertiary structure of the protein. Below the transition temperatures around 200 K, no jumps are observed. First short-lived jump attempts are seen at 210 K, at around 500 ps and 900 ps. At temperatures well above the transition range, such as 250 K and 300 K, conformational jumps on the side chains occur regularly. Such jumps mean moves of the side chain. This observation is in accordance with the results presented in section 4.2.3: After the dynamical transition, protein-water interaction reaches high levels, whereas it is negligibly small at temperatures below the transition. The movements of side chains, which start in the range of transition temperatures, are directly connected to this onset of interaction.

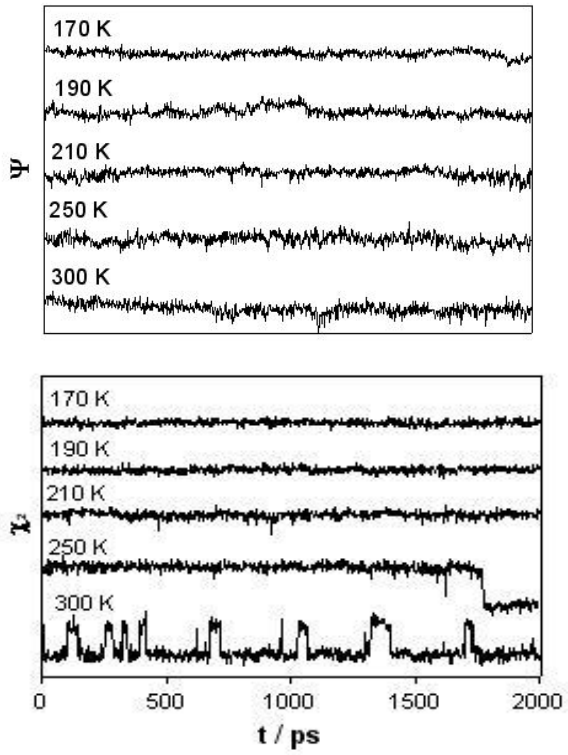

Figure 4.11 Torsional angle trajectories ψ (backbone) and χ_2 (side chain) for residue 75

4.3.2 Radial Distribution Functions

It has been shown that the protein-water interaction starting at the dynamical transition results in the high mobility of side chains at this temperature. The obtained data is not sufficient however to conclude if the side chains mobilize water molecules around them, or the increasingly mobile water molecules allow side chains to reach this level of mobility. Since these chains are the location of the total interaction, examination of processes including side chains and water molecules at specific sites can illuminate the basics of the dynamical transition. Labeling and tracking water molecules around specific residues shows that these water molecules switch places during the course of simulations at all temperatures, albeit with different rates, and other molecules from bulk water take their place with similar orientations. A better idea about this process can be presented by radial distribution functions. Figure 4.12 shows water radial distribution functions *g(r)* between O_{H_2O} and C_α of residue 75, the mobility of which was investigated in section 4.3.1. These functions, plotted by averaging over all timesteps, show the average density of water molecules at a distance *r* from the C_α of the residue. The graph is scaled by the average water density of the whole solution. In other words, *g(r)* is proportional to the probability of finding a water molecule in a shell with distance *r* to the residue C_α. At temperatures below the dynamical transition temperature, a first coordination shell forming in the range of 5.5 Å around the C_α atom, which has a maximum at 4.5 Å. A second coordination shell in the range 5.5 – 7.5 Å is also observed peaking at ca. 6.5 Å are seen. At longer distances, there seem to be another cluster of water molecules at 9 Å. This third shell is wider at 170 K, and it is a sharper peak at 150 K (data not shown). Although noise is present, the order of three groups of water at same distances is evident before the transition. Radial distribution function at 210 K and above still have the first coordination shell, but second and third shells are blended to bulk water. Though a hint of the second cluster is still present, it cannot be labeled as a coordination shell. The first shell also widens and starts to meld into bulk water part as temperature rises from 210 K to 300 K. The loss of order in water molecules around the side chain after 190 K can again be attributed to the dynamical transition and the start of sharp movements in the chain. It is only natural that the water molecules are not ordered into coordination shells as the side chain experiences discrete motions.

After the transition, the entire water component is affected by the protein as seen on Figures 4.8 and 4.9. As all side chains start to move, water molecules around them interact with these chains, and they interact with other waters. Also, since water molecules have a high mobility, molecules from the midst of bulk water travel near side chains and interact with them by replacing water molecules located there. This rotation also contributes to the fact that the entire bulk water feels the effect of moving side chains. To support these findings with a similar comparison to the functions given in Figure 4.12, radial distribution functions for bulk water are produced. Figure 4.13 consists of these functions. This time, the center atom is not the C_α of a residue, but O of a water molecule. In fact, the data was obtained by averaging radial distribution functions of 5 random bulk water molecules.

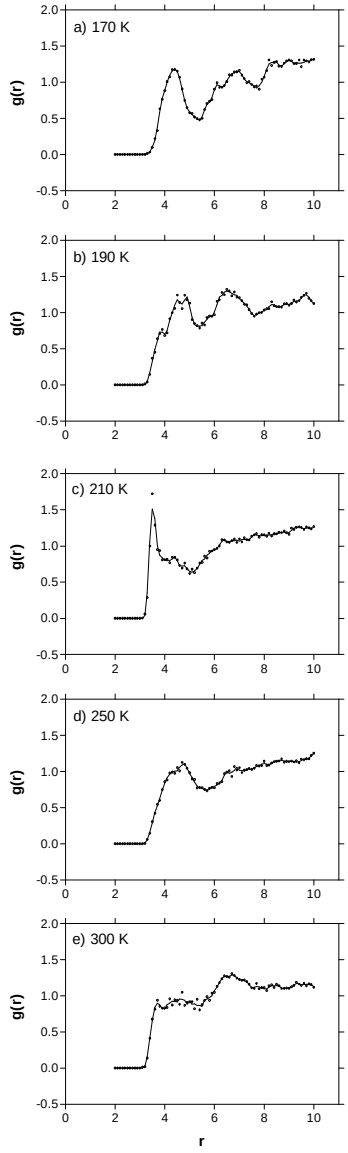

Figure 4.12 Water radial distribution functions around residue 75 at a) 170 K, b) 190 K, c) 210 K, d) 250 K and e) 300 K

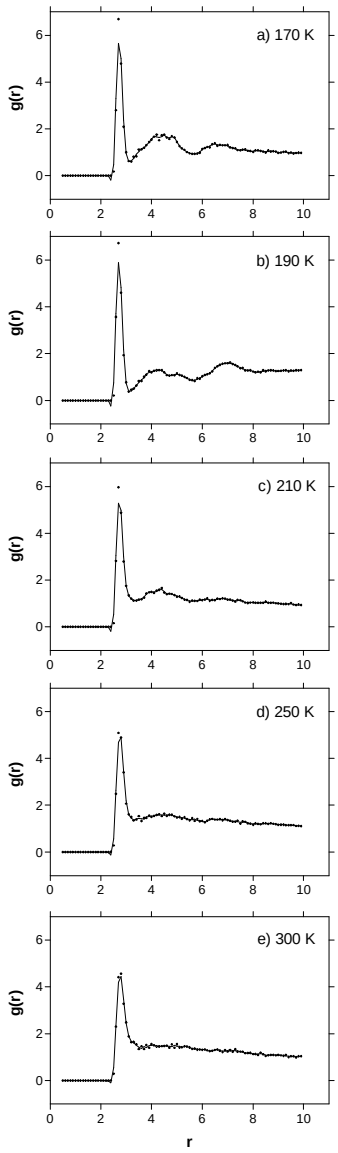

Figure 4.13 Bulk water radial distribution functions at a) 170 K, b) 190 K, c) 210 K, d) 250 K and e) 300 K

Bulk water radial distribution functions support the aforementioned conclusions. Functions below the transition temperature show a three coordination shell order, second and third of which completely disappear at higher temperatures. The first coordination shell never diminishes due to the ever-present short range order in water. The whole water component is experiencing the effect of the rotating and fluctuating side chains, as these observations agree.

5. CONCLUSIONS AND FUTURE WORK

In this work, dielectric properties of a protein solution consisting of lysozyme, water and chloride ions were calculated from MD simulations at different temperatures. At the temperature range studied, the protein remains folded, but the fluctuation behavior around that folded structure changes. In particular, at around 200 K, the protein goes through a dynamical transition or the “protein glass transition” and is functional only above this temperature, although it keeps the template overall structure well below this temperature. 2 ns long MD simulations in the range 150-300 were run with the exception of the 8 ns long 300 K simulation. The dielectric properties were used as analysis tools to investigate this *protein dynamical transition*, protein-water interaction and temperature dependence of this interaction, especially just below and right above the transition temperature. Mechanical analysis tools such as torsional angle trajectories and radial distribution functions of water molecules around the flexible protein surface residues were used to support the interpretation of the results provided by dielectric properties.

Auto- and cross-correlation functions of three components of the solution (protein, water, ions) were calculated. These correlations were fit to biexponential decay functions to enable analytical calculation of frequency dependent dielectric constants. A slow mode with a ca. 3000 ps relaxation time was found to be dominant in the protein dipole moment autocorrelation and a fast mode with ca. 7 ps was found to be dominant in the total water dipole moment autocorrelation.

Static and frequency dependent dielectric constants for protein and water components and the total solution were obtained for a range of temperatures using computations based on linear response theory. Static dielectric constants for lysozyme were found to fluctuate about 20 at different temperatures. This constant is close to those reported in literature for lysozyme and other proteins were found [8, 12, 14]. It

should be noted that the protein static dielectric constant does not change very much with temperature. If there is a slight trend, it remains below noise levels, so that it is impossible to capture. Water static dielectric constants showed a decay with increasing temperature as expected and they are smaller with respect to pure water molecules.

At temperatures below the dynamical transition, the frequency dependent water dielectric constants have a monotonic behavior. During the transition, this measurable shows a nick at the frequency of decay of the protein, which indicates that the entire water component is affected by the protein. This sign of protein-water interaction grows with increasing temperature. Cole-Cole plots were drawn for protein and water components and the total solution at different temperatures. These plots showed the same effect of interaction in the water component at and above 210 K, pointing to an onset of interaction at the dynamical transition temperature range.

Backbone and side chain torsional angle trajectories of several surface residues for the analyzed temperatures were obtained. While the backbone stays fluctuating about the same angular conformation, side chains begin to make conformational jumps after the transition. The start of interactions between protein and water molecules was tied to this emergence of enhanced mobility. Water molecules were labeled and monitored to find that different water molecules from bulk water replace the ones around the side chains continuously. It was observed that such a mobility exists even in the glassy regime, but the time scale of the diffusion process dramatically increases with the onset of the side-chain dihedral angle jumps.

As a more global quantification of the protein surface – water interactions, radial distribution functions of water molecules around the previously mentioned side chain residues and radial distribution functions of bulk water were plotted. Both sets showed a loss of order at the transition temperature. The model that we put forth thus suggests that the side chains gather sufficient energy to fluctuate vigorously enough to sample all the energy minima of the side chain torsional angles, while interacting with the immediate water around them; the surface water is also mobile enough to allow these conformational transitions. This effect then propagates to the entire bulk water by the continuous tumbling and diffusion of the surface water molecules into the bulk. The total volume of the water around the side chains in the simulations is very small

compared to the bulk, yet the effects of interaction with the protein are seen not only in these restricted regions, but in the entire water component.

The suggested model is involving protein-water interactions at side chains starting with the dynamical transition of the protein. Then the effect around side chains spreads to bulk water. Yet it must be stressed that the transition should not be seen solely as flowing from protein side chains to nearby water molecules to bulk water. In the absence of water molecules surrounding the side chains, it would not be possible for the side chains to gain the mobility they do at the transition. Without water molecules (or, in fact, any type of solvent that favorably interacts with their hydrophillic structure [30]) around them, the side chains would “stick” to the bulk of the protein instead of acquiring an open conformation that permits them to freely sample all the allowed conformational states. Thus, without the correct side-chain orientation, the mobility effects that propagate into the bulk water would not be observed. Water molecules do not only provide a medium for this transition, but are also actively a part of it. It is known that in absence of water the fluctuations in the torsion angle of side chains increase, but the dynamical transition is not observed [28, 29]. Therefore, it is possible to state that presence of water is necessary for this transition, but it is not sufficient. Temperature is also important both for protein fluctuations and mobility of water molecules around the side chains. As temperature increases, both fluctuations in the protein increase and water molecules become more mobile, allowing the side chains to experience rotational jumps. When the transition temperatures are reached, both effects are strong enough to start mobilization of the side chains. After that, the proposed mechanism of interaction spreading from side chains over near water molecules to the entire water component takes lead. It should also be stressed that in this study the interactions were analyzed by investigating electrical properties. Electrostatic interactions, especially for proteins in water, are long range interactions, the effects propagating into distances as large as 80 nm [54]. The propagation of the discussed dielectric effects to bulk water is therefore easier to track than it is for mechanical effects.

In future work, one possible study that would support these findings is to compute the dielectric relaxation behavior of a hypothetical protein-water system where the protein and the solvent are maintained at well-separated temperatures. Such simulations

were previously conducted by Vitkup et al. to study the fluctuations of the heavy atoms in the protein, a mechanical property [27]. Therein, the need for water to be at high enough temperatures was stated for the protein to gain large fluctuations neccessary for function. Tracking the actual water-protein interactions over large distances using the methodology of the current work is essential, however, to make conclusive statements about the overall functioning of the system.

REFERENCES

1. Gilson, M. K. & Honig, B. H. Calculation of electrostatic potentials in an enzyme active site. Nature 330, 84-86 (1987).

2. Fröhlich. Theory of dielectrics (Clarendon Press, Oxford, 1958).

3. Böttcher, C. J. F., Belle, O. C. V., Bordewijk, P. & Rip, A. Theory of electric polarization, 2nd ed. (Elsevier, Amsterdam, 1973), Amsterdam, 1973).

4. Grant, E. H., Sheppard, R. J. & South, G. P. Dielectric behaviour of biological macromolecules in solution (Clarendon Press, Oxford, 1978).

5. Nakamura, H., Sakamoto, T. & Wada, A. A theoretical study of the dielectric constant of protein. Protein Engineering 2, 177-183 (1988).

6. King, G., Lee, F. S. & Warshel, A. Microscopic simulations of macroscopic dielectric constants of solvated proteins. Journal of Chemical Physics 95, 4366-4377 (1991).

7. Simonson, T., Perahia, D. & Brünger, A. T. Microscopic theory of the dielectric properties of proteins. Biophysical Journal 59, 670-690 (1991).

8. Smith, P. E., Brunne, R. M., Mark, A. E. & Van Gunsteren, W. F. Dielectric properties of trypsin inhibitor and lysozyme calculated from molecular dynamics simulations. Journal of physical chemistry 97, 2009-2014 (1993).

9. Antosiewicz, J., Mccammon, J. & Gilson, M. Prediction of ph-dependent properties of proteins. Journal of Molecular Biology 238, 415-436 (1994).

10. Simonson, T. & Perahia, D. Internal and interfacial dielectric-properties of cytochrome-c from molecular-dynamics in aqueous-solution. Proceedings of the

National Academy of Sciences of the United States of America 92, 1082-1086 (1995).

11. Yang, L., Weerasinghe, S., Smith, P. & Pettitt, B. Dielectric response of triplex dna in ionic solution from simulations. Biophysical Journal 69, 1519-1527 (1995).

12. Loffler, G., Schreiber, H. & Steinhauser, O. Calculation of the dielectric properties of a protein and its solvent: Theory and a case study. Journal of Molecular Biology 270, 520-534 (1997).

13. Boresch, S., Ringhofer, S., Hochtl, P. & Steinhauser, O. Towards a better description and understanding of biomolecular solvation. Biophysical Chemistry 78, 43-68 (1999).

14. Boresch, S., Hochtl, P. & Steinhauser, O. Studying the dielectric properties of a protein solution by computer simulation. Journal of Physical Chemistry B 104, 8743-8752 (2000).

15. Yang, L., Weerasinghe, S., Smith, P. E. & Pettitt, B. M. Dielectric response of triplex dna in ionic solution from simulations. Biophysical Journal 69, 1519-1527 (1995).

16. Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Jr., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J. & Kollman, P. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. Journal of the American Chemical Society 117, 5179-5197 (1995).

17. Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. Science 288, 1604-1607 (2000).

18. Daniel, R. M., Smith, J. C., Ferrand, M., Hery, S., Dunn, R. & Finney, J. L. Enzyme activity below the dynamical transition at 220 k. Biophys. J. 75, 2504-2507 (1998).

19. Fenimore, P. W., Frauenfelder, H., Mcmahon, B. H. & Parak, F. G. Slaving: Solvent fluctuations dominate protein dynamics and functions. Proc. Natl. Acad. Sci. USA 99, 16047-16051 (2002).

20. Pal, S. K., Peon, J. & Zewail, A. H. Biological water at the protein surface: Dynamical solvation probed directly with femtosecond resolution. Proc. Natl. Acad. Sci. 99, 1763-1768 (2002).

21. Tsai, A. M., Udovic, T. J. & Neumann, D. A. The inverse relationship between protein dynamics and thermal stability. Biophys. J. 81, 2339-2343 (2001).

22. Baysal, C. & Atilgan, A. R. Relaxation kinetics and the glassiness of proteins: The case of bovine pancreatic trypsin inhibitor. Biophys. J. 83, 699-705 (2002).

23. Baysal, C. & Atilgan, A. R. Relaxation kinetics and the glassiness of proteins: Coupling of time scales. Biophys. J. 88, 1570-1576 (2005).

24. Dvorsky, R., Sevcik, J., Caves, L. S. D., Hubbard, R. E. & Verma, C. S. Temperature effects on protein motions: A molecular dynamics study of rnase-sa. J. Phys. Chem. 104, 10387-10397 (2000).

25. Tarek, M. & Tobias, D. J. Role of protein-water hydrogen bond dynamics in the protein dynamical transition. Phys. Rev. Lett. 88, 138101 (2002).

26. Tournier, A. L., Xu, J. & Smith, J. C. Translational hydration water dynamics drives the protein glass transition. Biophys. J. 85, 1871-1875 (2003).

27. Vitkup, D., Ringe, D., Petsko, G. A. & Karplus, M. Solvent mobility and the protein 'glass' transition. Nat. Struct. Biol. 7, 34-38 (2000).

28. Arcangeli, C., Bizzarri, A. R. & Cannistraro, S. Role of interfacial water in the molecular dynamics simulated dynamical transition of plastocyanin. Chem. Phys. Lett. 291, 7-14 (1998).

29. Lehnert, U., Réat, V., Weik, M., Zaccaii, G. & Pfister, C. Thermal motions in bacteriorhodopsin at different hydration levels studied by neutron scattering: Correlation with kinetics and light-induced conformational changes. Biophys. J. 75, 1945-1952 (1998).

30. Tsai, A. M., Neumann, D. A. & Bell, L. N. Molecular dynamics of solid-state lysozyme as affected by glycerol and water: A neutron scattering study. Biophys. J. 79, 2728-2732 (2000).

31. Caliskan, G., Kisliuk, A., Tsai, A. M., Soles, C. L. & Sokolov, A. P. Protein dynamics in viscous solvents. J. Chem. Phys. 118, 4230-4236 (2003).

32. Zaccai, G. Proteins as nano-machines: Dynamics-function relations studied by neutron scattering. J. Phys.: Condens. Matter 15, S1673-S1682 (2003).

33. Baysal, C. & Atilgan, A. R. Relaxation kinetics and the glassiness of native proteins: Coupling of timescales. Biophys. J. 88, 1570-1576 (2005).

34. Baloğlu, Ç. Refolding Kinetics of Lysozyme: Nuclear Magnetic Resonance & Molecular Dynamics Study of Hen Egg White Lysozyme. Ms. Thesis (Sabancı University, Istanbul, 2005).

35. Okan, O. B. Glassiness and Coupling of Timescales in Functional Proteins. Ms. Thesis (Sabancı University, Istanbul, 2005).

36. Karplus, M. & Petsko, G. A. Molecular dynamics simulations in biology. Nature 347, 631-639 (1990).

37. Leach, A. R. Molecular modelling: Principles and applications (Addison Wesley Longman, Singapore, 1996).

38. Brooks, R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. Charmm: A program for macromolecular energy, minimization and dynamics calculations. Journal of Computational Chemistry 4, 187-217 (1983).

39. Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. Namd2: Greater scalability for parallel molecular dynamics. Journal of Computational Physics 151, 283-312 (1999).

40. Haile, J. M. Molecular dynamic simulations: Elementary methods (John Wiley & Sons, New York, 1992).

41. http://hyperphysics.phy-astr.gsu.edu/hbase/electric/dielec.html. (2006).

42. Kirkwood, J. G. The dielectric polarization of polar liquids. The Journal of Chemical Physics 7, 911-919 (1939).

43. Onsager, L. Electric moments of molecules in liquid media. Journal of the American Chemical Society 58, 1486-1493 (1936).

44. Neumann, M., Steinhauser, O. & Pawley, G. S. Consistent calculation of the static and frequency-dependent dielectric constant in computer simulations. Molecular Physics 52, 97-113 (1984).

45. Neumann, M. Dielectric relaxation in water. Computer simulations with the tip4p potential. Journal of Chemical Physics 85, 1567-1580 (1986).

46. Neumann, M. Computer simulation and the dielectric constant at finite wavelength. Molecular Physics 57, 97-121 (1986).

47. Boresch, S. & Steinhauser, O. Presumed versus real artifacts of the ewald summation technique: The importance of dielectric boundary conditions. Berichte Der Bunsen-Gesellschaft-Physical Chemistry Chemical Physics 101, 1019-1029 (1997).

48. Kohler, F. The liquid state (Verlag Chemie, Weinheim, 1972).

49. Neumann, M. & Steinhauser, O. On the calculation of the frequency-dependent dielectric constant in computer simulations. Chemical Physics Letters 102, 508-513 (1983).

50. Neumann, M. & Steinhauser, O. On the calculation of the dielectric constant using the ewald-kornfeld tensor. Chemical Physics Letters 95, 417-422 (1983).

51. Caillol, J. M., Levesque, D. & Weis, J. J. Theoretical calculation of ionic solution properties. Journal of Chemical Physics 85, 6645-6657 (1986).

52. Caillol, J. M., Levesque, D. & Weis, J. J. Electrical properties of polarizable ionic solutions. I. Theoretical aspects. Journal of Chemical Physics 91, 5544-5554 (1989).

53. Diamond, R. Real-space refinement of the structure of hen egg-white lysozyme. Journal of Molecular Biology 82, 371 (1974).

54. Israelachvili, J. Intermolecular & surface forces, 2nd ed. (Elsevier, Oxford, 1991).