# VICINAL SOLVENT AND PROTEIN COUPLING MODULATES THE DYNAMICS OF PROTEINS

by
Ayşe Özlem Sezerman

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

Sabancı University
Spring 2007

# VICINAL SOLVENT AND PROTEIN COUPLING MODULATES THE SOLVENT DYNAMICS

A.Özlem Aykut

FENS, M.Sc. Thesis ,2007

Thesis Supervisor: Prof.Dr. Canan Atılgan

Thesis Co-Supervisor: Doç.Dr. O. Uğur Sezerman

## Abstract

The dynamics of a folded protein is studied in water and glycerol at a series of temperatures below and above their respective dynamical transition. The system is modeled in two distinct states whereby the protein is decoupled from the bulk solvent at low temperatures, and communicates with it through a vicinal layer at physiological temperatures. A linear viscoelastic model elucidates the less-than-expected increase in the relaxation times observed in the backbone dynamics of the protein. The model further explains the increase in the flexibility of the protein once the transition takes place and the differences in the flexibility under the different solvent environments. Coupling between the vicinal layer and the protein fluctuations is necessary to interpret these observations. The vicinal layer is postulated to form once a threshold for the volumetric fluctuations in the protein to accommodate solvents of different sizes is reached. Compensation of entropic-energetic contributions from the protein-coupled vicinal layer quantifies the scaling of the dynamical transition temperatures in various solvents. The protein adapts different conformational routes for organizing the required coupling to a specific solvent, which is achieved by adjusting the amount of conformational jumps in the surface-group dihedrals.

# PROTEİN ve PROTEİN YÜZEYİNİ SARAN ÇÖZGENİN BAĞLAŞMASININ ÇÖZGEN DİNAMİĞİNE ETKİSİ

A.Özlem Aykut

## Özet

Su ve gliserol içine konmuş katlanmış proteinin, dinamik geçiş sıcaklığı üstünde ve altındaki sıcaklıklarda dinamiği çalışıldı. Protein ve çözgen bir system olarak düşünülmüştür ve bu system iki durum için modellenmiştir. Birinci durumda protein ve çözgen düşük sıcaklıklarda birbirleri ile ilişki halinde değildir, ikinci durum ise fizyolojik sıcaklıklar icin gecerlidir ve bu durumda protein ve çözgen birbirleri ile bağlaşım içindedir. Bu bağlaşımı sağlayan ise protein yüzeyini saran çözgen (vicinal solvent) katmanıdır.Proteinin çatısal dinamiği incelendi. Protein çatısının gevşeme zamanlarının beklenenden düşük olması ise lineer viskoelastik bir model kullanılarak açıklandı.Bu model aynı zamanda, proteinin dinamik geçişi gerçekleştirdikten sora kazandığı esnekliğin ve farklı çözgenler içinde gösterdiği değişik esneklik hareketlerinin sebebini de açıklamaktadır. Bu gözlemler protein ve etrafını saran çözgen katmanının salınım etkileşimleri ile gerçekleşmektedir. Protein etrafını saran çözgen katmanının, değişik boyutlardaki çözgenlerin protein katalitik bölgesinde yer alabilmesi için proteinin hacimsel dalgalanmalarına bir eşik oluşturduğu varsayılmıştır. Değişik solventlerde görülen geçiş sıcaklıklarının birbirlerine gore oranlanabilmesi ise protein ve etkileşimde olduğu çözgen katmanından gelen etropik ve enerjetik kompansasyon ile olmaktdır.

Protein yuzeyde bulunan yan grouplarin değişik dihedral konformasyonlara geçmesi ile kendi etrafındaki çözgen katmanı ile etkileşime geçer.

# ACKNOWLEDGEMENTS

It is not possible to fully express my gratitude and feelings to several people whom I give more credit in completing my thesis than myself. In this little section I will attempt to my best.

I would like to thank Canan Atılgan for being the fountain of inspiration and encouragement since I got to know her. She opened me the gates of a scientific career for me and showed me how to continue by believing in what you are doing to gain success. Moreover, I am deeply affected from her pursuit of perfection albeit I failed to comply most of the time.

I would also like to thank Ali Rana Atılgan for his sparkling and fruitful ideas. He has been an excellent supporter through out this thesis study by always exposing me to his bombardment of new ideas and challenging questions. I will never forget the importance of springs in life which I have learned from him.

I would not be in this place if my co-advisor O. Ugur Sezerman did not encourage me. He always had intelligent and creative ideas. He is the one who had been a model of a successful scientific researcher for me since my childhood. He gave me the chance of working in a Tubitak project at Sabancı University and opened the gate of doing research.

Deniz Turgut has been a great officemate and a trusted supporter. I can not imagine working and finishing this thesis without his help. I always admire his intelligence and perfectionism. I would like to thank Alimet Sema Özen for her support when I was a beginner at computational work. Aslı Nalbant, Canan Barıştıran, Burcu Saner has been great friends. They were always there when I needed. Also I am grateful to Canan Barıştıran for her great help in editing my thesis. Özgür Gül, Günseli Bayram Akçapınar, Gökhan Kaçar and Murat Mulayim have been great labmates. And many thanks to Çetin Baloğlu for his support with NAMD.

I am most thankful to Tolga Aykut. His never-ending love and support was a safety net whenever I feel lost. His smile made the lights shine when I was in dark. And finally a million thanks to the best parents and family I have. My part in this work is dedicated to them.

# Table of Contents

# List of Figures

# List of Tables

*"Certainly no subject or field is making more progress on so many fronts at the present moment than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms."*

*Feynman Lectures on Physics,1988*

# 1. Introduction

Proteins execute and control essentially all functions in living organisms and they do it so elegantly and efficiently, with designs pined by billions of years of evolution. They are the most versatile macromolecules which can function as catalysts, transport and store other molecules such as oxygen, provide mechanical support and immune protection, generate movement, transmit nerve impulses, and they control growth and differentiation.

A large variety of processes in living organisms significantly depends on protein activity. In many of those only the structure of a protein determines its function ($\alpha$-keratin in hair), protein dynamics is important to many others. All biological processes which involve protein motion are based on the protein dynamics. Muscle contraction, for example, is based on the combined action of actin and myosin. The ability to change the conformation is also essential for the function of many transport proteins, proteins involved in signal transduction, proteins in the immune system, and numerous enzymes [1]. In many enzymes, conformational changes serve to enclose the substrate, thereby preventing its release from the protein and ideally positioning it for the protein to perform its function. Dynamics play an important role not only in the native state of many proteins where they are functional, but also in the mechanism by which a protein reaches that native conformation, the protein folding process.

Proteins are dynamic systems with the internal motions playing an essential role in their functionality [2]. Several experiments demonstrated the importance of internal motions in biomolecular function, including the hinge bending modes for opening and closing active sites, the flexibility of tRNA, the fluctuations required for ligand entrance and exit in heme proteins and the role of configurational entropy in proteins and nucleic acids [3-7]. All of the functions of proteins require flexibility because they need at least small rearrangements of atoms in response to those events. Proteins, like polymers, display such flexibility over a vast range of time scales and distances, covering motions which scale from femtoseconds (individual bond vibrations) to picoseconds (small group fluctuations) to nano- and microseconds and longer (collective motions of groups of bonded and non-bonded atoms). The corresponding length scales are from fractions of an Angstrom to tens of nanometers [8].

The large part of the current knowledge about proteins is derived from experimental data, especially X-ray crystallography and Nuclear Magnetic Resonance (NMR) [9]. There is currently no experimental technique that allows monitoring of protein conformational changes

at atomic resolution as a function of time at time-scales of nanoseconds. Computer simulation techniques provide the ultimate detail concerning individual particle motions in the picosecond to microsecond time range [2].

All of the organisms have an optimum temperature at which they live and function. Since the protein flexibility is essential for function, and then one might expect that at this optimum temperature any protein from that organism will display a flexibility that is sufficient for its function. This is exactly what is observed[8]. Temperature therefore can be a potential probe for the functional role of thermally driven motions for proteins. Yet, changing the temperature affects all motions at the same time, and this probe would have been more valuable if it could be used to perturb selectively one class of motions. The experiments show that not all the motions in proteins were sensitive to temperature in the same way.

The earliest evidence for unexpected temperature dependence of protein dynamics came from a study in which the atomic motions of the iron atom in the heme group of sperm whale metmyoglobin were measured using Mossbauer spectroscopy [10]. Biphasic behavior was observed with a low temperature regime whose linear behavior resembled those of small molecule compounds and a higher temperature region, which displayed a linear steep temperature dependence that could not be described with reference to model compounds. When the high temperature regime of the curve is extrapolated to very low temperatures, a negative fluctuation value is expected at absolute zero which suggests a hysterisis in the system. The situation is resolved by a transition leading to the terminology "unexpected" or "unusual" for the temperature dependence.

Similar phenomena were observed with the experiments using different techniques. Alben et al. studied the infrared spectra of bound and photodissociated states of carboxy-myoglobin (Mb-CO) from 5.2 K to 300 K [11]. In the bound state of Mb-CO, the known lines have center frequencies, widths and absorbances that are independent of temperature between 5.2 and 160 K. Above 160 K, one of these gradually shifts. The low-temperature photodissociated state (Mb) shows three lines, the absorbances of which depend on temperature. One was tentatively assigned to free CO in the heme pocket and the others to CO weakly bound to the heme or heme pocket wall, respectively. This observation lead to an explanation that the difference between Mb and deoxyMb is not due to an interaction of the flashed off ligand with the protein, but is caused by an incomplete relaxation of the protein structure at low temperatures.

Inelastic scattering experiments on myoglobin probe the motions of hydrogen atoms in the protein on the 0.1-100 ps time scale [12]. Elastic scattering intensity could be modeled by

harmonic behavior below approximately 200 K. above 200 K, motions occur that could not be modeled as harmonic motions of individual atoms. The extra mobility above 200 K was reflected in a strong temperature dependence of the mean square atomic displacements. This is correlated with the phenomenon previously observed specifically for heme iron by Mossbauer spectroscopy [10]. This was also observed with flash photolysis experiments and correlated with the transition in the hydration shell of myoglobin and with the temperature dependence of ligand binding rates at the heme iron [13].

Infrared spectroscopy experiments monitored the height of the energy barrier to CO rebinding in photolyzed CO myoglobin [14]. CO could be flashed off, but rebinds quickly below 180 K. Above approximately 200 K, CO begins to escape from the protein to the solvent. This implies that the fluctuations required for rebinding or the release of CO from the binding pocket is quenched below 200 K.

Hoffman and co-workers conducted experiments which showed the similar behavior at the same temperature [15]. The electron transfer rates between two proteins were shown to go to zero at approximately 200 K. The interpretation was that the complex formation becomes impossible below 200 K.

Parak and Petsko examined the X-Ray structure of myoglobin at 80 K. They extracted the atomic dynamic parameters from temperature-dependent diffraction data [16]. It was found that the overall structure at 80 K is similar to that at 300 K except that the volume is smaller. Comparison of Debye-Waller parameters at 80 K with the values obtained earlier at 250 - 300 K indicated that the protein at 80 K is more rigid. The mean square displacements vs. temperature plots for the backbone showed a discontinuity of the slope at approximately 200 K. The magnitudes of the mean square atomic displacements at 80 K suggested that intermolecular motions are frozen out. However some atoms have large magnitude mean square displacements at 80 K which was interpreted as providing evidence for conformational substates.

This behavior is not unique to myoglobin. Other experiment done with ribonuclease-A at nine different temperatures ranging from 98 to 320 K showed an expansion with increasing temperature (0.4% per 100 K) and this expansion is linear [17]. Also Debye-Waller factors exhibited biphasic behavior, with a small positive slope at low temperatures and a larger positive slope at higher temperatures. The discontinuity occurred at 200 - 220 K, a characteristic temperature found for myoglobin.

Molecular dynamics (MD) simulations provided further insight into the nature of the transition. MD was used to probe the atomic motions of Mb-CO with a hydration shell as a

function of temperature [18]. At low temperatures there is one fast decay process, and at high temperatures there is an additional slow relaxation process. The average atomic fluctuations show that protein behaves harmonically at low temperatures. At approximately 200 K, a transition in atomic fluctuations was seen. Above this temperature atomic motions displayed both harmonic and anharmonic behavior. Heavy atom dihedral transitions were monitored as a function of temperature. Dihedral transitions concerning backbone atoms occur only above the transition temperature. In another study, molecular dynamic simulations were performed during which the dihedral transitions are prohibited [19]. In that case, dehydrated Mb-CO exhibited only harmonic fluctuations where as hydrated Mb-CO exhibited both harmonic and anharmonic motions.

All of these results suggest that at low temperatures there is a purely vibrational motion with one fast decay process where the molecule is trapped in localized regions of the conformational space. At higher temperatures, there is more anharmonic motion with fast and slower relaxation processes occurring simultaneously, which arise from the combination of vibrations within the wells (substates) and transition between them [18, 20].

Molecules interact with one another by interactions – enzymes with their substrates, hormones with their receptors, antibodies with their antigens. The strength and specificity of interactions are highly dependent on the medium in which they take place, and the majority of biological interactions take place in water. Thus, proteins exist in a watery environment; their surfaces are covered with a layer of bound solvent, typically 2-3 ordered water molecules per residue [8]. This water can be considered as an intrinsic part of the protein structure compared to bulk solvent. All of the observations of a transition in protein dynamics lead to question; Is the transition due to a transition in the bulk solvent, the bound solvent, the protein alone or the combination of them ?

Calorimetry and IR experiments of the water of hydration in myoglobin crystals and solutions at subzero temperatures revealed the involvement of the bound water in the dynamic transition [12]. It is found that the hydration water remains amorphous at low temperatures. Moreover, a broad transition between 180 K and 270 K is found depending on the hydration level.

Detailed analysis of protein crystal structures, especially of the same protein in multiple cosolvent systems, indicates that bound water molecules on protein surfaces can be divided into several distinct classes. Some are completely disordered and though invisible to X-Ray experiments and can be observed by NMR. A second class of bound water molecules are observed in many, but not all structures of the same protein and have positions and

interactions that can differ slightly. They may have one or more unsatisfied potential hydrogen bonding interactions. Finally, there are bound water molecules that are always observed in the same positions, making the same interactions in nearly all structures of the same protein [8].

Analysis made on the hydrogen bonds by probing the shift of the vibrational peak frequencies with temperature of myoglobin in sucrose/water and glycerol/water solutions revealed more information [21]. At the transition, a characteristic change in the temperature slope of the O-H stretching mode was observed, which correlates with the discontinuity of the thermal expansion coefficient. The temperature-difference spectra of amide bands show the same tendency, indicating that stronger hydrogen bonding in the bulk affects the main chain solvent interactions in parallel. Also at lower temperatures it has been observed that there is an increased strength in the protein-water bonds. Moreover, the hydrogen bond network contracts which leads to frozen-in structures. This effect may represent the molecular mechanism underlying the dynamic transition observed for the mean square displacements of the protein atoms.

Inelastic neutron scattering in the low frequency range at full and low hydration levels showed that at room temperature hydrated samples exhibit more pronounced quasi elastic spectrum due to diffusive motions than the dehydrated samples [22]. Furthermore, the analysis of the corresponding line shapes suggests that water modifies mainly the amplitude but the characteristic time of fast protein motions. In contrast, at low temperatures, dehydrated samples show larger motional amplitudes than the hydrated ones. Water coupled librations of the polar side chains are depressed in the hydrated system due to strong intermolecular hydrogen bonding.

MD simulations by Vitkup et al. showed that the magnitudes of the protein fluctuations are largely determined by solvent viscosity [23]. As the temperature is increased above 180 K, the solvent gains mobility and determines the atomic fluctuations although intrinsic protein effects become important at lower temperatures. As a result, high solvent viscosity plays an essential role in inhibiting the dynamics of the protein below transition [24]. The rate of conformational changes in myoglobin measured with nanosecond lasers hinted that depending on the solvent viscosity, the dynamics is controlled by the solvent (at high solvent viscosity), the protein (at low solvent viscosity), or a crossover regime where both are effective [25].

The biological function of proteins is affected by the structural fluctuations among the conformational sub-states [26] and the solvent plays a key role in their activation [27]. It has been reported that there is a correlation between protein dynamics and the thermal motion of

water [28]. The water molecules promote these fluctuations through a hydrogen bond network, so that the conformational fluctuations in the protein occur on the same (picosecond) time scale as the fluctuations in water [29]. MD simulations on lysozyme showed that the dynamics of solvent – protein interactions along the surface control the structural relaxation of the protein as a whole [24], the effect propagating into the core of the protein through intermolecular interactions, also for the glycerol solvent. The protein, in return, has an influence on the solvent dynamics so that the solvent molecules in close proximity of the protein surface show decreased mean square fluctuations and increased relaxation time. It is known that the shell of water around the protein surface display dynamics that is distinct from the bulk. This layer – termed biological water or vicinal water – has slower relaxation behavior measured by femtosecond lasers [30] and MD simulations [31], displaying a bimodal character on the pico- and sub-picosecond time scales. Both relaxation times are slower by about one order of magnitude in the vicinal layer, the effect being most effective up to 4 Å from the surface. This maybe due to the hydrogen bonding or other close-range favorable interactions between the solvent molecules and the side chains on the surface of the protein; in fact, at a distance of ca. 7 Å all water molecules have bulk properties. The hydrogen bonding slows down the rotational and translational motion of solvent molecules in close proximity to protein surface [32], so that the solvent dynamics are suppressed as the solvent molecules get closer to the protein.

The motion of the solvent molecules is slow and traps the protein molecules in long lived conformations below the transition; conversely, above the transition solvent dynamics is fast enough to let the protein sample several conformations [23]. MD simulations by Vitkup et al. showed that the magnitudes of the protein fluctuations are largely determined by solvent viscosity [23].

In another work the temperature dependent thermodynamical and dynamical properties of the protein dynamical transition was investigated through MD simulations on bovine pancreatic trypsin inhibitor [33, 34]. They have characterized the  transition, and proposed that with the onset of the transition, an insertion of time scales occurs that operate between the slow mode motions of the protein and fast vibrational dynamics [34]. Thus, the dynamics between these motions that are well separated in frequency becomes coupled, leading to a functional protein. Therein, it was suggested that those intermediate time scales were brought about by the side-chain conformational jumps.

In another study, MD simulations of myoglobin surrounded by a shell of water were performed by using a dual heat bath method [35]. This method allows the protein and solvent

to be held at different temperatures. Therein, the mean square fluctuations of protein non-hydrogen atoms were investigated and it was shown that the protein dynamical transition when hydrated is driven by temperature dependent changes in the solvent dynamics. This idea was supported by the finding that holding the solvent at high or low temperatures abolishes the dynamical transition, whereas varying the solvent temperature with the protein fixed at 300 K recovers it.

The dynamics and function of proteins are found to be coupled to motions in the bulk solvent and the hydration shell [36]. Three types of protein motions were identified: Solvent-slaved, hydration-shell-coupled, and vibrational (non-slaved). The motions involving large-scale conformational changes are solvent-slaved motions and follow the α-fluctuations in the bulk solvent. They govern such processes as the entrance and the exit of ligands. Hydration-shell coupled motions follow fast β-fluctuations in the hydration shell, and most likely involve side chains and permit processes such as the passage of ligands inside the protein [32, 36]

Neutron scattering analysis of xylanase in powder form, in $D_2O$ and in four two-component perdenaturated single-phase cryosolvents showed the general features of the dynamic transition behavior; the fast picosecond fluctuations of concentrated protein solutions follow those of the solvents [37]. Also, it has been shown that the solutions of proteins in different cryosolvents (different proportions of the mixed solvents), display similar picosecond dynamic transition behavior.

An understanding of the role of solvent on protein dynamics will inevitably lead to an understanding of how proteins function while sensing different environments. Also it will give insights for designing protein-solvent formulations with improved shelf lifes.

All of the results presented above indicate that the dynamics of the protein are strongly influenced by the solvent, and therefore the temperature dependence of the solvent. The question remains as to the mechanisms utilized by the solvent so as to determine the dynamics of the folded protein below and above the transition. We present here the results of a series of extensive MD simulations on *Trichoderma reesei* Endoglucanase III (1H8V) in two different solvents, water and glycerol. Proteins function in a watery environment and a dehydrated protein is functionless. On the other hand, glycerol is a solvent which is mostly used as cryo-protectant at low temperatures. When the proteins or enzyme formulations are kept in glycerol, they do not degrade for long periods of time, keeping their functionality. A lot of research is being made about the mechanism lying under this protection. In this study, glycerol has been chosen to bring a new understanding to this phenomenon and the data

obtained from protein in water will be used as a baseline to compare results with that of glycerol.

We first characterize the common features of the dynamical transition with thermodynamical quantities and thermal fluctuations in the systems. We then use viscoelastic and thermodynamical models to elucidate the findings. We employ the side-chain torsional angle trajectories as evidence to mark the differences in the two solvent systems.

# 2. Theoretical Background

Computer simulations came into fashion among scientist when hardware became affordable. Molecular simulations are about solving the statistical mechanical equations by numerical techniques [38] . Simulations can provide the ultimate detail concerning individual particle motions as function of time. Thus, they can be used to address specific questions about the properties of a model system; often more easily than experiments, on the actual system, provided that the properties of interest are on the right time- and length-scale. This chapter gives a general overview of the molecular dynamics theory and molecular dynamics simulation method employed in this study.

## 2.1 Molecular Dynamics

MD simulations employ Newton's laws of motion to simulate many-body systems under force fields. The force fields provide approximations for the forces on each atom [38]. The MD trajectory is obtained by solving the differential equations embodied in Newton's second law (F = ma):

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \qquad (\ 2\text{-}1\ )$$

This equation describes the motion of a particle of mass $m_i$ along one coordinate ($x_i$) with $F_{x_i}$ being the force on the particle in that direction. The force acting on particle $i$ can also be described as the gradient of the potential energy, $V$:

$$F_i = \left[ -\nabla U(r_1, r_2, \ldots, r_n) \right]_i = -\frac{\partial U(r_1, r_2, \ldots, r_n)}{\partial r_i} \qquad (\ 2\text{-}2\ )$$

The first MD simulation performed by Alder and Wainwright in 1957 [39]. They have performed the MD simulation of a condensed phase system using a hard-sphere model and did not employ the continuous force potentials [38]. In a more realistic model, the force on each particle will change whenever the particle changes its position, or whenever any of the other particles with which interacts changes its position. The first MD simulation using

continuous potentials was performed by Rahman on argon [40]. Introduction of continuous potentials give rise to a many-body problem that cannot be solved analytically. The equations of motions are integrated using finite difference method. All algorithms for integrating the equations of motion assume that the position and dynamic properties can be approximated with Taylor series expansions.

Verlet algorithm is the adapted scheme for the MD simulations reported in this study. It uses the positions and acceleration at time $t$ , and the positions from the previous step, $x(t-\Delta t)$, to calculate new positions at $t+\Delta t$, $x( t+\Delta t )$ as shown in the equations below;

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

Velocities and third order terms cancel each other out leaving accelerations and 4th order error terms behind. This implies that the Verlet algorithm which is apparently 2nd order is more precisely a 4th   order algorithm [41]. The ability to solve x in both directions in time brings out the feature that the Verlet algorithm is suitable for modeling conservative systems. Combining this with the Newton's second law, one ends up with:

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

Velocities may still be estimated with various methods one of which produces an estimate at half-step forward:

$$v(t + 1/2\Delta t) = \frac{x(t + \Delta t) - x(t)}{\Delta t}$$
( 2-5 )

On the other hand, a modified version, the so-called velocity-Verlet [42] algorithm, explicitly calculates the velocities as well as the positions and accelerations simultaneously. This is achieved through a three-step computation which involves;

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2$$

$$v(t + \Delta t) = v(t) + \frac{1}{2}a(t)\Delta t + \frac{1}{2}a(t + \Delta t)$$

( 2-6 )

For biomolecules, the potential energy corresponds to a set of semi-empirical functions that model the effective interactions between the particles. There has been a considerable effort to develop reliable interaction functions or force fields for proteins. Typical examples are AMBER [43], CHARMM [44], OPLS [45] and GROMOS [46]. In general, these force fields depend on the functional forms and parameters used to model each interaction. The interactions are modeled by two types of physical terms: Non-bonded interactions (charge-charge and van der Waals interactions) and bonded interactions (stretching, bending, and torsions). In addition, non-physical terms to restrain and constrain the system may also be included [1].

## 2.1.1 Non-bonded Interactions

The non-bonded interactions are divided into electrostatistic and dispersion/repulsion (induced dipole, Pauli repulsion) contributions. Therefore, each non-bonded interaction can be modeled using a Coulombic potential and Lennard-Jones (LJ) potential [47-50].

$$E_{non-bonded}(i, j) = E_{LJ}(i, j) + E_{Coulomb}(i, j)$$

$$E_{LJ}(i, j) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

( 2-7 )

$$E_{Coulomb}(i, j) = \frac{1}{4\pi\varepsilon_o}\frac{q_i q_j}{r_{ij}}$$

where partial charges are $q_i$ and $q_j$ , $\varepsilon_o$ is the vacuum permittivity, the Lennard-Jones parameters are $\sigma_{ij}$ (LJ-radius) and $\varepsilon_{ij}$ (LJ-energy) (shown in 1 with a and b, respectively). The collision diameter $\sigma_{ij}$ is the separation distance at which the energy is zero. $\varepsilon_{ij}$ represents the depth of the energy well at the minimum atomic distance.
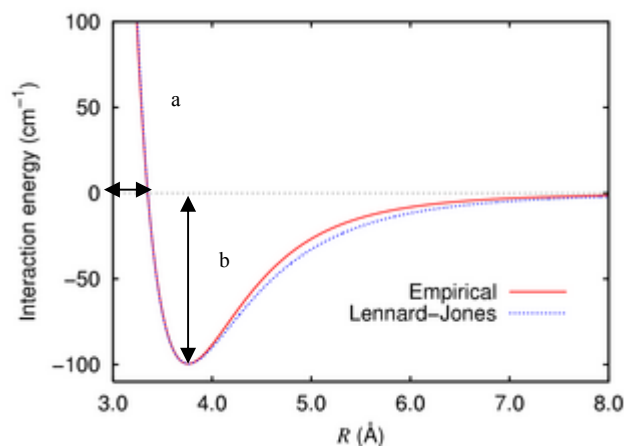
**Figure 2-1. The Lennard-Jones Potential**

The most time consuming part of the molecular dynamics simulation (or indeed, of an energy minimization) is the calculation of the non-bonded energies and /or forces [38]. The numbers of bond stretching, angle bending and torsional terms in a force field model are all proportional to the number of atoms but the number of non-bonded terms that need to be evaluated increases as the square of the number of atoms (for a pairwise model) and thus of order $N^2$ generally. In this study, particle mesh ewald sum method is employed to calculate long range interactions and the terms increases with $N\log(N)$. In principle, the non-bonded interactions are calculated between every pair of atoms in the system. The Lennard-Jones potential falls off very rapidly with distance: at 2.5σ the Lennard-Jones potential has just 1% of its value at σ. This reflects the $r^{-6}$ distance dependence of the dispersion interaction. The most popular way to deal with the non-bonded interactions is to use a non-bonded cutoff and to apply the minimum image convention. In the minimum image convention, each atom 'sees' at most just one image of every other atom in the system, which is repeated infinitely with periodic boundary conditions (for details see section 2.1.4 below). The energy/force is calculated with the closest atom or image. When a cutoff is employed, the interactions between all pairs of atoms that are farther apart than the cutoff value are set to zero, taking into account the closest image.

## 2.1.2    Bonded Interactions

The bonded terms $E_{bonded}$ are the sum of at least three terms. These are, bond stretching $E_{bond}$ between two covalently bonded atoms:

$$E_{bond} = \sum_{bonds} K_r (r - r_{eq})^2 \qquad\qquad \textbf{( 2-8 )}$$

where $r$ is the bond length between bonded atoms, and the parameters $K_r$ and $r_{eq}$ are defined for each type of pair atoms, such that the former represents the spring constant, and the latter is the average bond length.

Bond-angle bending, $E_{angle}$, term ensures that the bond angle fluctuates around its equilibrium value $\theta_{eq}$:

$$E_{angle} = \sum_{angles} K_\theta (\theta - \theta_{eq})^2 \qquad (\text{ 2-9 })$$

where $\theta$ is the angle between the atom triplet $i$-$j$-$k$, with atoms $i$-$j$ and $j$-$k$ being covalently bonded. $K_\theta$ controls the strength of the fluctuations around $\theta_{eq}$ and they are defined for each type of atom triplets.

The dihedral-angle term, $E_{dihedral}$ controls the rotations around the torsional angles and the barrier crossings between its $n$ minima:

$$E_{dihedral} = \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] \qquad (\text{ 2-10 })$$

to set the interactions for the quadrupole of atoms $i$-$j$-$k$-$l$. The angle $\psi$ is defined as the angle between planes $i$-$j$-$k$ and $j$-$k$-$l$. $V_n$, $\gamma$ and $n$ are set for each type of atom quadruplets.

## 2.1.3    Constraint Dynamics

The constraint dynamics is handled by the RATTLE algorithm. The RATTLE method allows one to consider atomic connectivity without harmonic bonds [38]. Valence bonds vibrate at high frequency and impose a small integration time-step to a simulation. To circumvent this problem, one may constrain the bond length so as to use a larger time step. The equality is usually written down in the form of a holomonic constraint:

$$r_k^2 - a_k^2 = 0 \qquad (\text{ 2-11 })$$

where $a$ fixes the bond length to the value of interest. In a constrained system, the coordinates of the particles are not independent of each other, and the equations of motion in each of the coordinate directions are related. The constraint forces lie along the bonds at all times. For each constrained bond, there are two equal, but opposite forces on the two atoms that comprise the bond, and therefore no energy is put into the system.

### 2.1.4 Periodic Boundary Conditions

Periodic boundary conditions enable a simulation to be performed using a relatively small number of particles, in such a way that the particles experience forces as they were in bulk fluid [38]. There are several possible shapes for space filling unit cells. Some , like a rhombic dodecahedron and the truncated octahedron  [38] are closer to a sphere than a cube and are therefore more economical for studying an approximately spherical macromolecule in solution, since fewer solvent molecules are required to fill the box given a minimum distance between macromolecular images.

In order to establish periodic boundary conditions for a cubic simulation box, it is replicated in all the space to from an infinite lattice. During the simulation, when a molecule leaves a simulation box, its image enters from the opposite side of the box, thus keeping the density of the box constant. The simulation box forms a convenient axis system for measuring the coordinates of the *N* molecules. To avoid self interaction, the box has to be larger than two times the cutoff distance used for non-bonded interactions [38, 47-50].

### 2.1.5 Constant Temperature Control with Langevin Dynamics

Langevin dynamics is a means of controlling the kinetic energy of the system, and thus controlling the system temperature [38, 47-50]. The method uses the Langevin equation for a single particle:

$$m_i \frac{d^2 x_i(t)}{dt^2} = F_i[x_i(t)] - \gamma_i \frac{dx_i(t)}{dt} m_i + R_i(t) \qquad ( \text{ 2-12 })$$

Here, two additional terms on the right hand side accompany the ordinary force acting on the particle. The second term represents a frictional damping that is applied to the particle with frictional coefficient $\gamma_i m_i$. The third term represents the random forces which act on the particle as result of the solvent interaction. These two terms are used to maintain particle kinetic energy to keep system temperature at a constant temperature.

## 2.2 Molecular Dynamics Simulation Protocol

A series of MD simulations under constant temperature – volume conditions were performed on Endoglucanase enzyme (Protein Data Bank [51] (PDB) code: 1H8V) in water and glycerol at temperatures spanning the range below and above the transition. 1H8V is a

218 amino acid protein (24.5 kDa), neutral pI, glycoside hydrolase family 12 cellulase that lacks a cellulose-binding module. It has a large substrate binding groove formed by the β sheets in which the active site residues are located (Figure 2). Active site residues Asp99, Glu116 and Glu200 from a carboxylic acid trio [52]. The main focus is given to these residues to understand the dynamic behavior of the catalytic site and the groove before and after the dynamical transition.
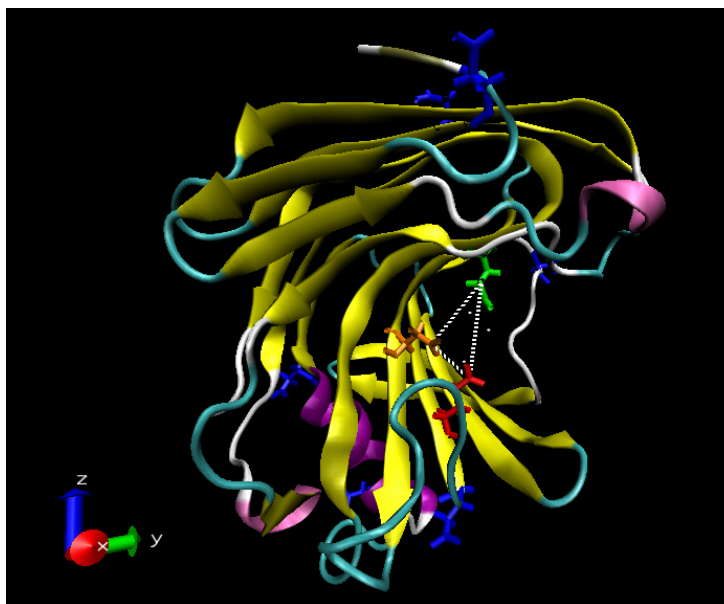


**Figure 2-2. The three-dimensional structure of the protein *Trichoderma reesei* Endoglucanase III (PDB code: 1H8V). The three catalytic residues are highlighted in color : ASP 99 (orange), GLU 116 (red), GLU 200 (orange). The refst of the Asp residues studied are also shown in blue.**

The NAMD package was used to model endoglucanase III, water (TIP3P) and glycerol [53]. The force field parameters were used from CharmM 22 for the former two and CharmM 27 for the latter [44]. The electrostatic potentials on glycerol atom surfaces were calculated using the software package GAUSSIAN 03 [54]. The equilibrated glycerol box containing 400 molecules was formed and optimized with the [55]; the equilibrated box of TIP3P water presented in the NAMD package is used for solvation.

The protein-solvent mixtures were formed by using VMD 1.8.5 solvent plug-in version 1.2 [56]. The protein was soaked in a solvent box such that there is at least a 5 Å layer of solvent in each direction from any atom of the protein. Then the protein-solvent mixtures were neutralized using VMD auto-ionize plug-in version 1.2. The protein-water system contains two sodium ions and one chloride ion along with 3055 water molecules, whereas the

protein-glycerol system neutralized with one sodium ion has 592 glycerol molecules. The constructed protein-water box has 12301 atoms with box dimensionality 111.8×51.3×64.2 Å. The protein-glycerol system was formed of 11452 atoms with dimensions of 111.8×54.9×66.9 Å.

Long range electrostatic interactions were calculated using particle mesh Ewald (PME) method. The cutoff distance for non-bonded van der Waals interactions was set to 12 Å with a switching function cutoff of 10 Å. Rattle algorithm was used to fix the bond lengths to their average values. During the simulations, periodic boundary conditions were used and the equations of motion were integrated using the Verlet algorithm with a step size of 2 fs. The temperature was maintained at the desired value during the equilibration stages using temperature rescaling and during the data collection stages using Langevin dynamics. The latter treats each atom separately, balancing a small friction term with Gaussian noise to control the temperature.

The protein-solvent mixtures were minimized until the gradient tolerance was less than $10^{-2}$ with the conjugate gradients algorithm implemented in NAMD. The protein – solvent mixture was pre-equilibrated at the simulation temperature and at constant volume for 500 ps. Then the data collection run was performed at constant volume for 2 ns. The coordinate sets were saved every 2 ps intervals for subsequent analysis. The temperature range is 120 – 300 K and 130 – 385 K for the protein in water and glycerol, respectively. The temperature is incremented with steps of 10 – 15 K; the last structure saved from the lower temperature run is used as input for the equilibration run at the next temperature. A total of 34 runs were performed for these systems.

# 3. Results and Discussion

## 3.1 The dynamical transition in water and glycerol

There has been a lot of research for understanding the glassy relaxation behavior of proteins [27] and RNA chains [57]. Recent incoherent neutron scattering measurements of elastic intensities of proteins [58] identify the conformational flexibility as essential for enzyme catalysis and that the positional fluctuations are important for the functionality because those positional fluctuations behave like a lubricant [59].

In this study we seek to understand the effect of two different solvents – water and glycerol – on the overall thermodynamical and dynamical behavior of the system, where the term "system" refers to the protein and the solvent as a whole. We therefore first seek the dynamical transition temperature of this protein in these two solvents. The phenomenon of dynamical phase transitions in proteins has first manifested itself in neutron scattering studies. Those studies are significant, since they clearly show the effect of water as solvent in dampening of the molecular flexibility and fluctuations below a temperature regime of 190-220 K [59-61]. When the mean squared fluctuations are plotted against temperature one observes a change in the slope when the protein is hydrated in water. However, for completely dehydrated powders, such a transition is absent.

Fluctuations in the thermodynamical entities such as the energy or volume are used to depict phase transitions. Their characterization is achieved by monitoring the relevant susceptibilities, e.g. the heat capacity or isothermal compressibility. Heat capacity at constant volume, $c_v$, is computed from the trajectories by using the relationship $\langle (E - \langle E \rangle)^2 \big/ k_B T^2 \rangle$ [38]. In these calculations, the trajectory at each temperature is divided into 10 chunks of 200 ps length, each with 100 data points. The temperature dependence of heat capacity for the two systems is displayed in Figure 3-1 and the uncertainty in the data is shown with the error bars. Note that the uncertainty in the results from the water – protein systems is larger in general. Nevertheless, it is possible to predict a transition temperature of 164 ± 4 K and 277 ± 5 K for 1h8v in water and glycerol, respectively.
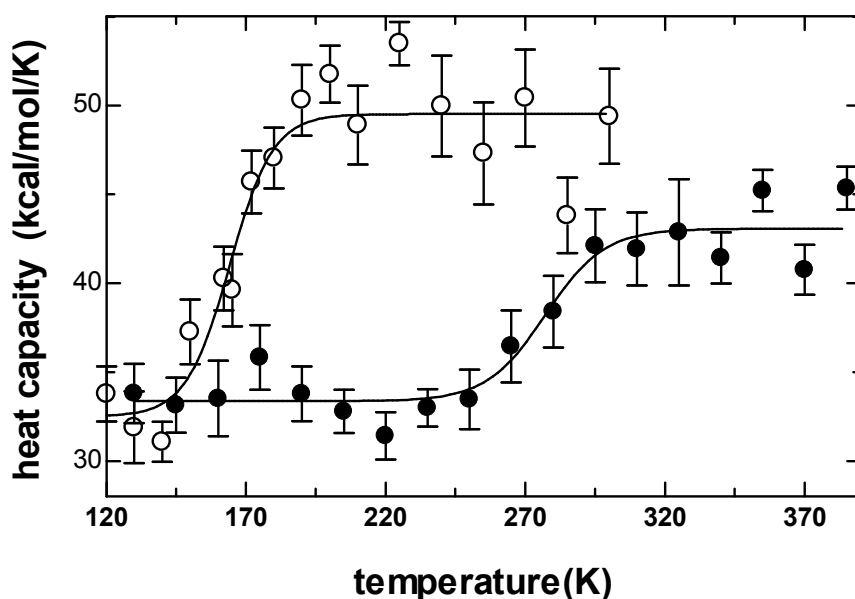
**Figure 3-1. Temperature dependence of signature system properties monitored for detecting the dynamical transition: System heat capacity. Results from protein – water and protein – glycerol systems are shown with open and filled circles, respectively, the best fitting lines are to Boltzmann sigmoidal functions see ref [34].**

Neutron scattering measurements monitor the average fluctuations in the hydrogen atoms of the protein, $<\mathbf{u}^2>$. The theory used in interpreting those results factors out the tumbling motions undergone by the protein. In the current study and in our previous work [33, 34], following a similar treatment, we monitor the average fluctuations in $C_\alpha$ atoms of the protein, defined as $<\Delta\mathbf{R}^2>$. To achieve monitoring the internal motions only, i.e. to factor as much of the overall tumbling motion of the protein as possible, we apply the following protocol: For any 200 ps piece of the trajectory, we first make a best-fit superposition of the recorded structures to the initial structure by minimizing the root mean square deviations of the $C_\alpha$ atoms. Then the average structure, $<\mathbf{R}(T)>$, from the 100 best-fitted structures is computed. Here the brackets denote the time average. Finally, another best-fit superposition of the recorded structures to this average structure is made. Each structure of this final trajectory is denoted by $\mathbf{R}(t,T)$ and the coordinates of the $i^{\text{th}}$ residue is given by $\mathbf{R}_i(t,T)$. The fluctuation vector for a given residue $i$ at a given time $t$ from a given trajectory obtained at temperature $T$, $\Delta\mathbf{R}_i(t,T)$, is thus the difference between the position vectors for the $i^{\text{th}}$ residue of the best-fitted and the average structures,

$$\Delta\mathbf{R}_i(t,T) = \mathbf{R}_i(t,T) - \langle\mathbf{R}_i(T)\rangle \qquad\qquad (\text{ 3-1 })$$

The mean square fluctuations are then obtained from an average over both time and residue index as $\langle \Delta \mathbf{R}^2 \rangle = \langle \overline{\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i} \rangle$. Note that it is not possible to completely separate the internal motions from the external using this protocol.

In figure 3-2, examples are shown for the $C_\alpha$ atoms at temperatures below, during and above the transition, where the transition temperatures are depicted from heat capacity analysis. The details of the residue-residue fluctuations are very similar, although the magnitude of the fluctuations differs markedly. The common feature of these systems is the average structure of the protein about which the fluctuations occur, an equilibrium property. This is a coarse approach to the analysis of the wealth of data produced by MD and will not give sufficient information on the details of the phase transition.

a)                                    b)
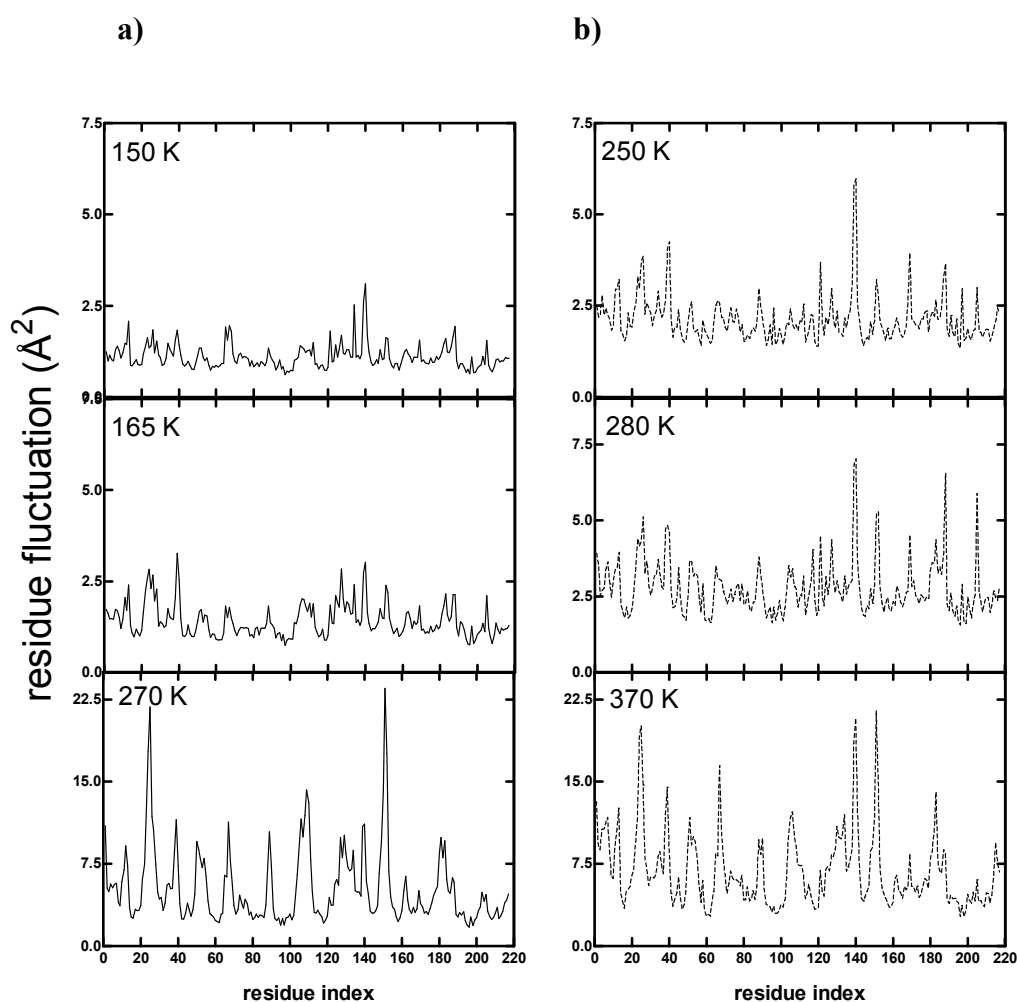


**Figure 3-2. a) Fluctuations of $C_\alpha$ in water below the transition (150 K), during the transition (165 K), above the transition (270 K). b) Fluctuations of $C_\alpha$ in glycerol below the transition (250 K), during the transition (280 K), above the transition (370 K).**

It is also possible to depict the dynamical transition from an analysis of residue fluctuations. As a sample of protein – solvent mixture is heated starting from a temperature well below the dynamical transition, the fluctuations are expected to follow a curve that may mainly be depicted to be linear which extrapolates to zero fluctuation at absolute zero as would be expected thermodynamically. The dynamical transition is defined to occur at the temperature where a deviation from this linear behavior of $\left\langle \Delta R^2 \right\rangle$ versus T is captured. Nevertheless, using chunks of different sizes, 200 ps – 10 chunks, and 400 ps – five chunks, it was verified (data not shown) that the nature of the transition is captured, and the transition temperature predicted does not change [34]. The new data again fall on a straight line, albeit the y-intercept not extrapolating to absolute zero anymore, but to a negative fluctuation that suggests a hysterisis in the system [62]. Along with the onset of the transition, it is expected that a small fraction of the total population of particles experience barrier crossing events caused by large amplitude fluctuations of the protein conformations [58]. The results are shown in figure 3-3; transition temperature of 166 ± 6 K and 267 ± 3 K for 1h8v in water and glycerol, respectively are predicted.
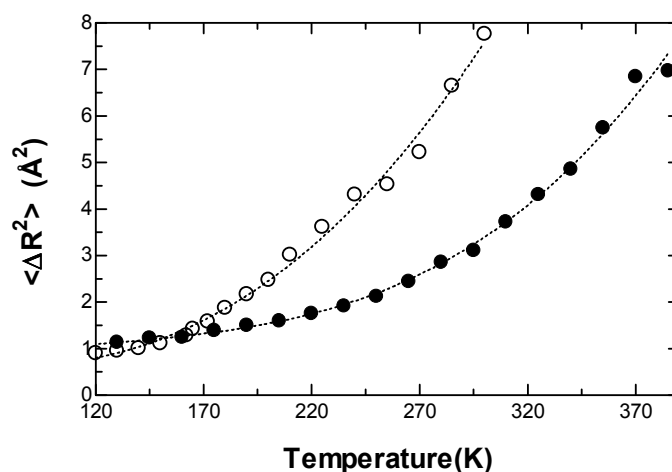


**Figure 3-3. Temperature dependence of fluctuations averaged over all residues protein – water and protein – glycerol systems are shown with open and filled circles, respectively. Dotted lines are shown to guide the eye.**

Whereas the magnitude of the fluctuations experienced by the backbone $C_\alpha$ atoms, $<\Delta R^2>$, gives an overall idea on the nature of the dynamical transition, the motion of the fluctuation vector carries a wealth of supplemental information. The relaxation of the $\Delta R$ vector can be characterized by a relaxation function of time and temperature, $C(\tau, T)$,

$$C(\tau,T) = \frac{\overline{\langle \Delta\mathbf{R}(0,T) \cdot \Delta\mathbf{R}(t,T)\rangle}}{\overline{\langle \Delta\mathbf{R}^2(T)\rangle}} \qquad (3\text{-}2)$$

where the bar and the brackets denote the average over all residues and the time average, respectively. $C(t,T)$ usually cannot be modeled with a single exponential decay, because there are many different contributing homogeneous processes with different relaxation times and the non-exponential behavior is shown in figure 3-4.
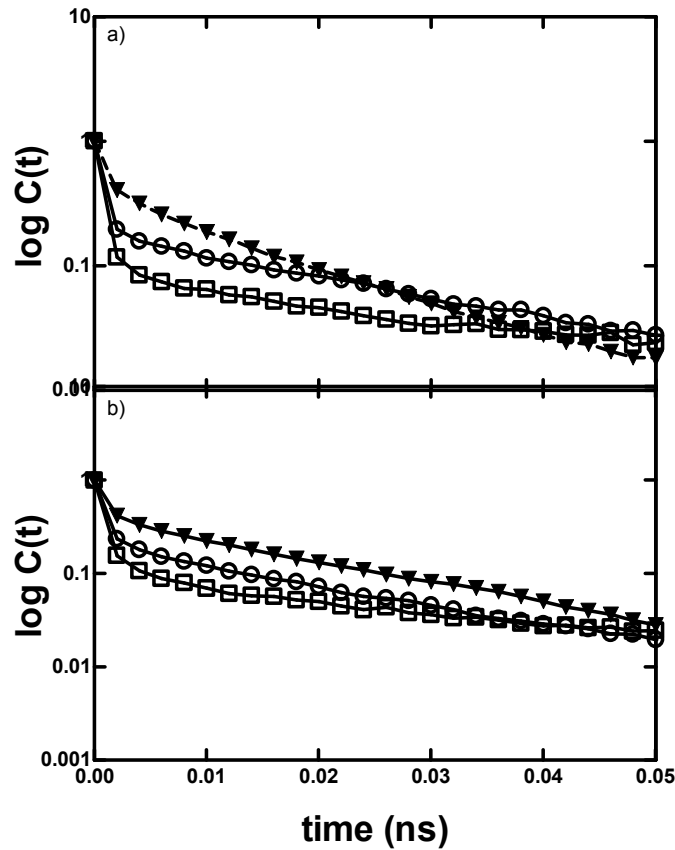


**Figure 3-4. a) Relaxation functions, $C(t,T)$ of $C_\alpha$ in water below the transition 150K-red; during the transition 165K-green; above the transition 270K-blue. b) Relaxation functions, $C(t,T)$ of $C_\alpha$ in glycerol below the transition 250K-red; during the transition 280K-green; above the transition 370K-blue.**

Contributions will result in heterogeneous dynamics and assuming all contributing processes show single exponential decay, each with relaxation time $t_i$, the relaxation function can be represented by [33] :

$$C(t,T) = \sum_{i=1}^{n} a_i \exp\left(\frac{-t}{\tau_i}\right)$$ ( 3-3 )

$a_i$ is the weight with which each of these processes contribute to the observed relaxation so that $\Sigma\, a_i = 1$. Sometimes it is possible to know the number and nature of all these processes; e.g. dielectric relaxation may usually be described by a biexponential fit. However, under many circumstances this is not possible, and furthermore all processes do not need to display Debye relaxations. One may resort to approximate the above equation by the Kohlrausch – Williams – Watts expression [63, 64];

$$C(t,T) = \exp(-t/\tau_e)^{\beta}$$ ( 3-4 )

As it can be seen from figure 3-3, at time scales faster than 20 ps there is a sharp decay that is not represented by the fit. This may be due to the power law behavior that is governing dynamics in this regime. In fact, in the sub-ps regime, non exponential contributions to relaxation phenomena are known to exist [27], but such time scales and motions are assumed to be averaged out at the time and length scales explored by the $C_\alpha$ fluctuation dynamics. In the present study we investigate the relaxation of positional fluctuations, which probes the collective dynamics which are at time scales slower than 20 ps.

In equation 3.4, the stretch exponent, $\beta$, is a quantity between 0 and 1, whereas $\tau_e$ is an equivalent relaxation time for the decay of the $C(t,T)$ function. Note that although a value of $\beta = 1$ usually implies a single process with a single exponential decay, departure from 1 does not necessarily involve more and more complicated dynamics. In a previous study, it was shown that two equivalently contributing processes with well separated relaxation times lead to a lower $\beta$ exponent than three such processes, where a third process with an intermediate time scale is inserted between the original ones [34]. The former is a "simpler" dynamics than the latter at the outset. As such, the $\beta$ exponent manifests how different contributing processes come together, leading to the observed relaxation behavior. The quantity $\beta$ was previously used to depict the dynamical transition [33], and its variation with the temperature in water and glycerol is shown in figure 3-5. The predicted transition temperature for 1h8v in water and glycerol is 182 ± 1 K and 280 ± 5 K, respectively.
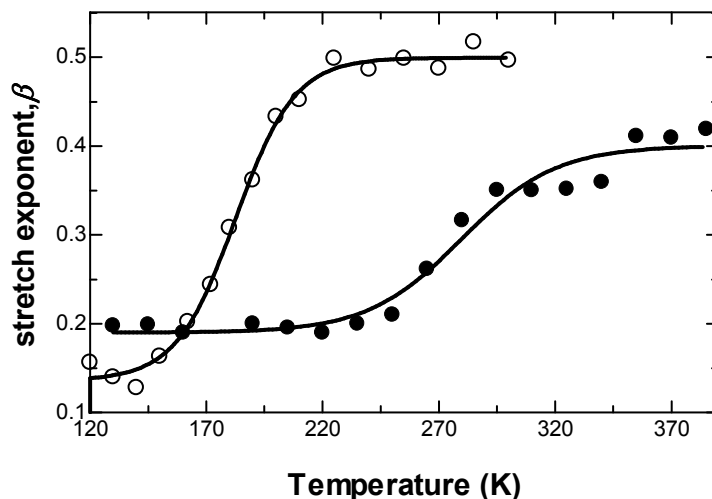
**Figure 3-5. The temperature dependence of the stretch exponent, protein – water and protein – glycerol systems are shown with open and filled circles, respectively, the best fitting lines are to Boltzman sigmoidal functions see ref [34].**

The estimated transition temperatures are summarized in Table 1 below. Note that the predictions using the three different methodologies are the same within error bars, whereas the physical origin of how the predictions are made varies: Heat capacity data belong to the overall system, including both the protein and the solvent, monitoring the energy fluctuations in the system, based on information of entropic nature. $\langle \Delta R^2 \rangle$, on the other hand, measures the amount of fluctuations experienced by the main chain $C_\alpha$ atoms. In this respect, it is a quantity that indirectly measures the consequences of the dynamics undergone by the system in different solvent – temperature conditions. Finally, $\beta$ is derived from the same $C_\alpha$ fluctuations; yet, instead of measuring the scalar amount of deviations from an average value, it stems from the dynamics of how the fluctuations vary in time. In particular, it recapitulates the combination of the many processes contributing to the relaxation of the monitored vector. In what follows, we will use how the dynamical transition surfaces in all these different phenomena to get a deeper understanding of the mechanism underlying the coupling between the dynamics of the folded protein and the surrounding solvent.

**Table 1. Protein Transition Temperatures Obtained by Various Approaches**

| | Protein-water system | Protein-glycerol system |
|---|---|---|
| Heat capacity, $c_v$ | $164 \pm 4$ K | $277 \pm 5$ K |
| Residue Fluctuations, $<\Delta \mathbf{R}^{2>}$ | $166 \pm 6$ K | $267 \pm 3$ K |
| Stretch exponent, $\beta$ | $182 \pm 1$ K | $280 \pm 5$ K |
| Average | $171 \pm 11$ K | $275 \pm 13$ K |

## 3.2 Solvent – protein communication leads to the dynamical transition

Below their transition temperatures, the heat capacity of both the protein – water and the protein – glycerol systems are in the same range with values of $32.5 \pm 1.7$ and $33.4 \pm 0.5$ kcal/mol/K, respectively (Figure 3). Therefore, prior to the onset of the dynamical transition, the energy fluctuations in the system are similar in magnitude, independent of the solvent. However, above the transition, the former system reaches a value of ca. $49.5 \pm 0.9$ kcal/mol/K whereas the latter attains a lower heat capacity of ca. $43.1 \pm 0.6$ kcal/mol/K. This observation hints that the protein and solvent begin to act as a unified and unique system at temperatures they are functional; conversely, one might conjecture that the protein behaves independently form the solvent, below the transition temperature. In fact, similar observations hold for the other two system properties that we monitored in figures 3, 5, 7 where the low temperature tails of the curves superimpose on each other.

Below we develop a viscoelastic model whereby we treat the protein and the solvent molecules that reside on its surface (vicinal solvent) as the system, which together are in a bath made of the rest of the solvent molecules. The equilibrium dynamics of the protein in the solvent is modeled by Brownian dynamics,

$$\mathbf{Z}\Delta\dot{\mathbf{R}} + \mathbf{K}\Delta\mathbf{R} = \mathbf{F} \qquad (3\text{-}5)$$

where F is the vector holding the random forces acting on the protein and the vicinal layer, and $\Delta$R gives the mean fluctuations in the system entities, $\Delta R T = [<\Delta R>_p \ <\Delta R>_v]$. Note that, in particular, $<\Delta R>_p$ is given by the average of eq. 3.1 over all residues and $<\Delta R>_v$ represents the average over all vicinal solvent molecules. The constant matrix **K** holds the spring

constants of each entity, and their coupling as $\mathbf{K} = \begin{bmatrix} k_p & k_{pv} \\ k_{vp} & k_v \end{bmatrix}$ where the cross-terms are taken

to be equal. $\mathbf{Z}$ describes the frictional environment, , $\mathbf{Z} = \begin{bmatrix} \zeta_p & \zeta_{pv} \\ \zeta_{vp} & \zeta_v \end{bmatrix}$ with the diagonal

entries representing the friction caused by the bulk solvent on the protein ($\zeta_p$) and the bulk solvent on the vicinal solvent ($\zeta_v$). The off-diagonal terms relate the friction imposed by the protein and vicinal solvent on each other, taken to be equal. Employing the fluctuation-dissipation theorem $\mathbf{FF}^T = 2k_B T \mathbf{Z}$, the time dependent auto-correlations of the fluctuations are obtained [65]

$$\langle \Delta \mathbf{R}_i(0) \cdot \Delta \mathbf{R}_i(t) \rangle = k_B T \left( e^{-t\mathbf{Z}^{-1}\mathbf{K}} \mathbf{K}^{-1} \right)_{ii} \tag{3-6}$$

Where $i = 1$ for the solvent, and 2 for the vicinal layer. From eq. 3.6, the instantaneous elastic response of the system is given at time $t = 0$ with $<\Delta \mathbf{R}^2>_p = k_B T \ (\mathbf{K}^{-1})_{11}$. At temperatures below that of the transition $T^*$, there are no terms for the vicinal solvent, and we deal with the 1×1 matrix, namely a scalar, for which we recover the harmonic solution for the protein, $<\Delta \mathbf{R}^2>_p = k_B T \ / \ k_p$. At high temperatures, the average fluctuations of the protein coupled to the vicinal solvent are given by the first term on the diagonal of the solution with $<\Delta \mathbf{R}^2>_p = k_B T \ / \ k_p'$ where the force constant acting on the system is no longer that of the protein alone, but is a multiple of it that also depends on the elastic motion of the protein in the vicinal layer. In summary,

$$k_p' = \begin{cases} k_p & T < T^* \\ k_p \left( 1 - \dfrac{k_{pv}^2}{k_p k_v} \right) & T > T^* \end{cases} \tag{3-7}$$

At sufficiently high temperatures, we thus get the solution with higher slopes in the fluctuation versus temperature curves, modified by the coupling between the vicinal solvent and protein. The harmonic approximation is the basis of the force constant measured by elastic neutron scattering experiments, whereby the inverse slope of the $<\Delta \mathbf{R}^2>$ versus temperature curves yields the spring constant [59, 66, 67]. Similarly, from the limiting slopes of the curves in figure 3-3 we calculate $k_p = 4$ N/m, whereas $k_p' = 0.7$ and 0.9 N/m for the coupled protein-water and protein-glycerol systems, respectively.

Furthermore, since the time dependent autocorrelations are defined by eq. 3.2, the relaxation time for the auto-correlation of the $i^{\text{th}}$ component is obtained by integrating $\mathbf{C}(t)$ for the $i^{\text{th}}$ term:

$$\tau_i = \int_0^\infty \frac{\langle \Delta\mathbf{R}_i(0)\cdot\Delta\mathbf{R}_i(t)\rangle}{\langle \Delta\mathbf{R}_i^2\rangle}\,dt = \frac{\left(\mathbf{K}^{-1}\mathbf{Z}\mathbf{K}^{-1}\right)_{ii}}{\left(\mathbf{K}^{-1}\right)_{ii}} \qquad (\text{3-8})$$

Due to the linear nature of the model, this expression predicts simple exponential decays for all the correlations. As before, below the transition temperature the matrices reduce to single term expressions due to the absence of the vicinal layer, and the average relaxation time of the protein is simply given by $\tau_p = \varsigma_p / k_p$. In contrast, that above the transition may be defined to be $\tau'_p = \varsigma'_p / k'_p$. The expression for the effective friction in the latter below and above the transition temperature is then:

$$\varsigma'_p = \begin{cases} \varsigma_p & T < T^* \\ \varsigma_p\left[1 - \dfrac{k_{pv}^2}{k_p k_v}\left(2\dfrac{\varsigma_{pv}/k_{pv}}{\varsigma_p/k_p} - \dfrac{\varsigma_v/k_v}{\varsigma_p/k_p}\right)\right] & T > T^* \end{cases} \qquad (\text{3-9})$$

Note that each of the ratios that appear in the bottom expression in eq. 3-9 is less than one. The term in square brackets gives the deviation from the protein friction coefficient. One might argue that the protein – vicinal solvent coupling will manifest itself in the stiffness and not in the damping. In that case, the second term as a whole must be much less than one. From eq. 3.7, we know that the $k_{pv}^2/\left(k_p k_v\right)$ term is smaller than, but on the order of one. Moreover, the last term is negligible compared to the second because, relative to the protein, the time scale of the vicinal solvent ($\varsigma_v/k_v$) is expected to be much less than that of the coupling between the protein and the solvent ($\varsigma_{pv}/k_{pv}$). Thus, if there is a less-than-expected increase in the effective friction coefficient, hence the relaxation time, this will mainly be due to the further damping caused by the vicinal solvent on the protein.

In figure 3-6, we display the characteristic times obtained from the relaxation of the $C_\alpha$ atom fluctuations in the MD trajectories (eq. 3.2). These data are obtained directly from the area enclosed by the $C(t)$ curves. Although the processes recorded do not display a simple exponential decay, as conceived by the theory (eq. 3-8), a relative scale for characteristic times is nevertheless obtained. Since the external motions of the protein are eliminated by the

structural best-fitting procedure described in the methods, the recorded times are typical of the average internal motions of the protein backbone.

As temperature is increased, and processes that involve the collectivity of larger number of atoms emerge, the relaxation time will be modified so as to yield longer times. Here we assume that the friction coefficient that arises due to the effect of bulk solvent on the protein, $\zeta_p$, remains roughly constant over the temperature range studied since protein may be considered to be a molten solid [68]. The same treatment was also made by Ansari et al. in explaining the crossover between the solvent- and protein-controlled kinetics governing the conformational changes following the photodissociation of carbon monoxide in myoglobin using nanosecond laser measurements [25]. Here, the fact that the relaxation times well-below the transition temperature are constant in figure 3-6 corroborates this assumption. We further find that the term $\zeta_{pv}$, due to the friction arising from the coupling between protein and vicinal water, is roughly independent of temperature, since the relaxation times well-above the transition are also constant. Then the leading term in eq.3-9 shows that the main adjustment in $\tau'_p$ is due to the altered flexibility of the protein, namely if $\zeta_{p'} \approx \zeta_p$ then $\tau_p / \tau_p = k_p / k_{p'}$. However, although the expected increase in $\zeta_p$ is by a factor of 5.7 and 4.4 for water and glycerol solvents, respectively, the observed increase in the relaxation time is only nearly by a factor of 2 in both solvents. Thus, the presence of the vicinal solvent molecules does not only cause an increase in the flexibility of the protein, but also modifies the effective friction that is felt by the protein.

Furthermore, relaxation of the protein maybe related to functionality of the protein. Protein and the substrate should move in a synchronized fashion to communicate with each other. In other words, if the protein has a catalytic site with an opening cleft, the cleft will be closed and substrate cannot enter in unless the relaxation time is on the order of the substrate movement.
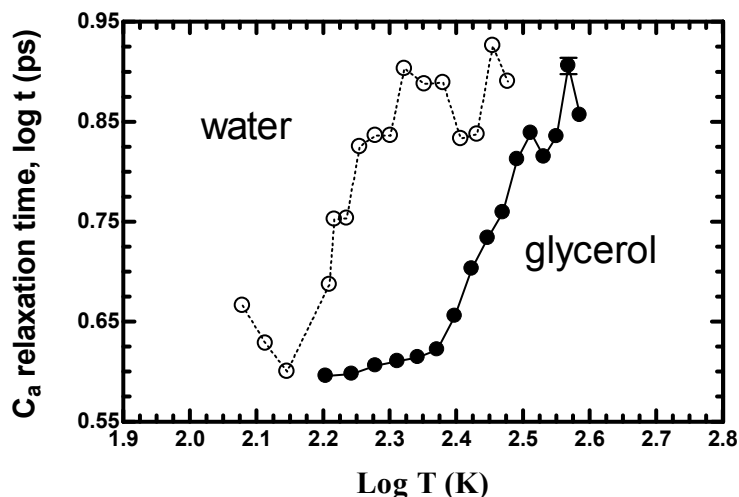
**Figure 3-6.** The $C_\alpha$ relaxation times; open circles: protein-water; filled circles: protein-glycerol systems.

## 3.3 The transition temperature is determined by the vicinal solvent layer

In figure 3-7, the data are plotted on the double logarithmic scale for both the relaxation time and the temperature, and the temperatures of water relaxation data are shifted by a factor $\log (T^*_{\text{glycerol}}/T^*_{\text{water}})$; two curves are superimposable. In other words, identical relaxation profiles are obtained if the temperatures are scaled by a factor $T^*_{\text{glycerol}}/T^*_{\text{water}}$. The main contribution in determining the location of the transition temperature may be estimated using the solvent-protein coupling ideas developed in the previous subsection and thermodynamic arguments.
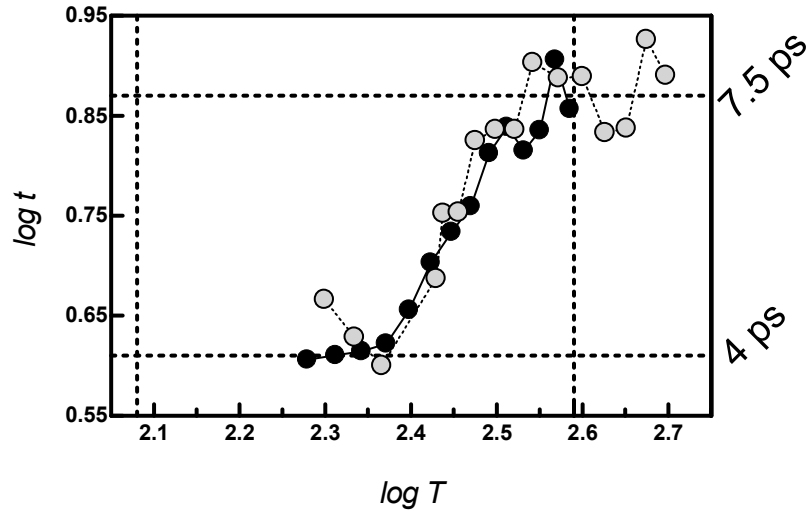
**Figure 3-7. The dotted curve shows the data for the protein – water system with the temperatures shifted by a factor of $T^{*}_{glycerol}/T^{*}_{water}$ (eq. 3-12).**

Suppose we idealize the system as having two stable states. In one, the protein is in direct contact with the bulk solvent (state A), and in the other, part of the solvent molecules organize into a vicinal solvent layer whereby the solvent molecules' motion is coupled to that of the protein, together existing in the bulk solvent (state B). These states are shown schematically in figure 3-3, upper panel. At temperatures exceeding the transition temperature, $T^{*}$, the coupled-solvent model (state B) is more stable, whereas at low temperatures the solute existing in the bulk solvent (state A) prevails.

The instantaneous change in the free energy of the system may be written as dF = –PdV – SdT, and under constant volume conditions, entropy at a given temperature is the slope of the F vs. T curve. Thus, thermodynamics require that F is a decreasing function of T, which at the bottom panel of figure 3-8, is schematically shown as a straight line. Stability requires that F remains a minimum, and we expect a phase transition between states A and B at T = T*, when the curve belonging to state B begins to remain below that of state A. At T = T* the free energies of the two states are equal:

$$\left.\begin{array}{c} F^{A} = F^{B} \\ U^{A} - T^{*}S^{A} = U^{B} - T^{*}S^{B} \end{array}\right\} \qquad \textbf{( 3-10 )}$$

We then write the contributions to the energy and entropy of the system, as the sum of those from the bulk solvent and the protein:

$$\left(U_p^A + U_{bulk}^A\right) - T^*\left(S_p^A + S_{bulk}^A\right) = \left(U_p^B + U_{bulk}^B\right) - T^*\left(S_p^B + S_{bulk}^B\right) \qquad \textbf{( 3-11 )}$$

We assume that the bulk of the solvent (denoted by the subscript *bulk*) does not feel the effect of the solvent molecules that contribute to the vicinal solvent layer that becomes coupled to the protein; $U_{bulk}^A = U_{bulk}^B$ and $S_{bulk}^A = S_{bulk}^B$. On the other hand, the difference between the energetic and the entropic contributions from the protein (in A) and protein coupled system (in B) is denoted $\Delta U_{p'} = U_p^B - U_p^A$ and $\Delta S_{p'} = S_p^B - S_p^A$ so that $T^* = \Delta U_{p'}/\Delta S_{p'}$. The subscript $p'$ refers to the fact that these differences take into account the interactions and microstates that develop with the organization of the vicinal solvent layer, additional to those intrinsic to the protein.
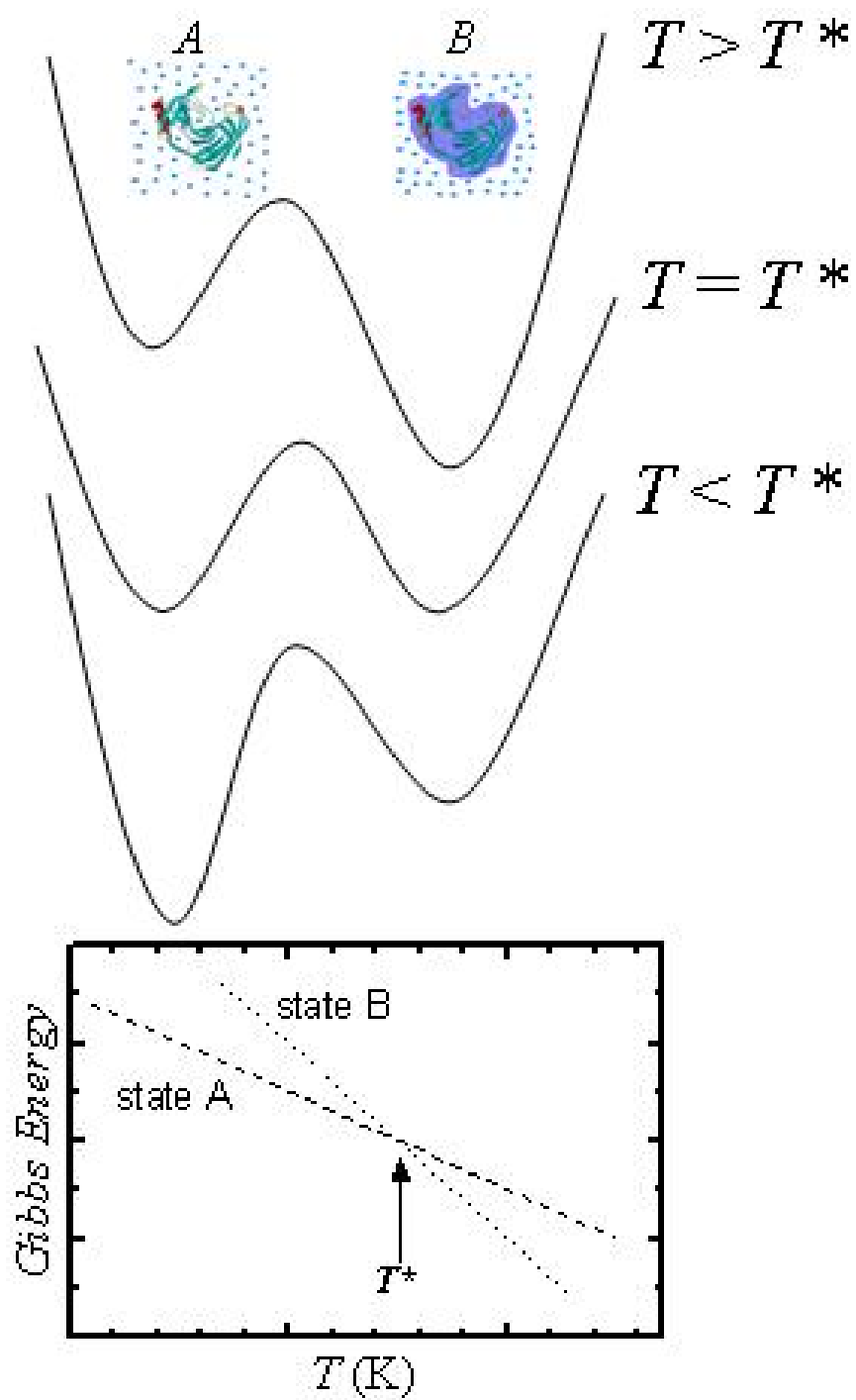
**Figure 3-8.** Scheme depicting the two-state model of vicinal water organization

Then the ratio of the transition temperatures in two different solvents, e.g. water and glycerol as in the current study, is given by the expression
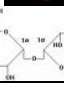
$$\frac{\left(T^*\right)_{glycerol}}{\left(T^*\right)_{water}} = \frac{\left(\Delta U_{p'}/\Delta S_{p'}\right)_{glycerol}}{\left(\Delta U_{p'}/\Delta S_{p'}\right)_{water}} \qquad (3\text{-}12)$$

The energetic contribution will predominantly be due to interfacial interactions in the additional layer of modified solvent that is organized around the protein. If there are $n$ solvent molecules in this layer, each with an accessible surface area of $a$, then $\Delta U \; \alpha \; (na)$. This organization also brings an entropic cost that is proportional to the size of this layer, which may be approximated by $(nv)$, where $v$ is the volume enclosed by the accessible surface of one solvent molecule, so that $\Delta S \; \alpha \; (nv)$. Assuming that the strength of the interactions in the interface is similar in these solvents, mainly due to the availability of –OH groups that interact favorably with the protein surface atoms along the surface of both types of solvents, the ratio of the transition temperatures may then be approximated by,

$$\frac{\left(T^*\right)_{glycerol}}{\left(T^*\right)_{water}} = \frac{\left(a/v\right)_{glycerol}}{\left(a/v\right)_{water}} \qquad (3\text{-}13)$$

We have computed the protein accessible surface areas and volumes of glycerol and water molecules by choosing the probe radius as 1.5 Å, an average value for the heavy atoms in the protein. Also, we have checked our model for another solvent which is trehalose and the probe radius is taken as 1.5 Å. Results are shown in Table 2.

**Table 2. Accessible surface area and volume of solvent molecules**

| Molecule | Surface area, $a$ (Å$^2$) | Volume, $v$ (Å$^3$) | $a/v$ (Å$^{-1}$)[†] | Molecular Structure |
|----------|---------------------------|---------------------|---------------------|---------------------|
| water | 31 | 16 | 1.9 |  |
| glycerol | 101 | 81 | 1.2 |  |
| trehalose[‡] | 268 ± 1 | 271 ± 6 | 1 |  |

[†]The $a/v$ ratios are the same for probe sizes ranging from 1.4 – 1.6 Å; results displayed here are for probe radius of 1.5 Å.
[‡]Values averaged over different conformers around the torsional angle connecting the rings

The ratio of the transition temperature of 1h8v in glycerol and water is 1.6, using the average values found in Table 1 and so is the prediction from eq.3-13. Similarly, the dynamic transition temperature of hen egg white lysozyme in water, glycerol, and trehalose were determined to be 190, 300 and 350 K, respectively [24, 69, 70]. Thus, glycerol again scales the transition temperature of lysozyme by a factor of 1.6. Note that the actual transition temperature in water is shifted by ca. 20 K due to the details of the protein structure. Similarly, trehalose scales $T^*$ by a factor of ca. 1.8 and the prediction from eq.3-13. is 1.9; this equation is still applicable, since the trehalose molecule also has many –OH groups along its surface, and is involved in atomic interactions with the protein similar to those of water and glycerol. Given the crudeness of this approach with the assumptions involved, the success of the predictions points to the fact that many contributing effects cancel in energy and entropy differences of eq.3-13 so that the scaling in the transition temperature is mainly due to the vicinal solvent effect.

## 3.4  Nature of solvent–protein communication through side–chain conformational transitions

The behavior of average atomic fluctuations ($<u^2>$ in neutron scattering experiments, $<\Delta R^2 >$ in this work) is consistent with the presence of a distribution of energy minima with successively higher barriers being crossed as the temperature is increased, the characteristic times for these events being on the ps scale [37]. The relaxation time of the protein in both water and glycerol above their respective transition temperatures are found to be ca. 7 ps, hinting that the dynamics of the protein backbone is independent of the solvent. Thus, the average contribution of all processes affecting the relaxation is similar in both solvents, whereas the effective mechanisms which contribute to relaxation are different, as we discuss below.

Tarek and Tobias have attributed the onset of the dynamical transition to the relaxation of the hydrogen bond network in the water via solvent translational displacements [32]. Therein, it was also shown that inhibition of these motions has been influential in suppressing protein motion, especially in the fluctuations of the side-chain atoms. Corroborating this view, in our previous work, we have shown that the communication between protein and solvent during and above the dynamical transition is manifested in the onset of conformational jumps of the torsional angles in surface group side-chains [34]. These jumps were put forth as the motions that provide the intermediate time scales responsible for the communication between the slow globular protein motions, and fast local dynamics.

We study the $\chi_2$ angles of the seven Asp residues in the protein, six of which are located on the surface and one belongs to the catalytic triad (figure 2). We record the fraction of angles that display at least one conformational jump at a given temperature during the observation time frame of 2 ns (figure 3-9).
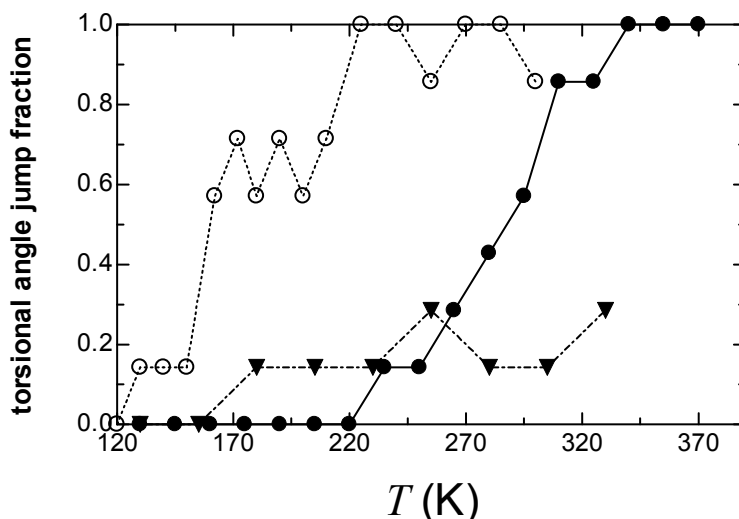


**Figure 3-9. The fraction of the $\chi_2$ torsional angles that display at least one jump between the conformational states at a given temperature for the seven Asp residues in the studied protein. Results from protein – water and protein – glycerol and protein-only (no solvent) systems are shown with open circles, filled circles and filled triangles, respectively.**

As anticipated, the tendency of the dihedrals to sample their rotameric states increases as the temperature is raised. However, the behavior is solvent dependent, as well as the exact location of the monitored Asp residue. For example, in glycerol no jumps are recorded in any of these residues below the transition; conversely, in water occasional jumps occur, albeit not at the same residue (active site residue Asp 99 at 130 K, flexible loop residues Asp5 at 140 K, and Asp126 at 150 K).

We check that direct interactions between solvent molecules and side-chain heavy atoms are necessary for the conformational jumps to occur by monitoring MD trajectories of Asp99 and Glu116 $\chi_2$ dihedrals of 1h8v obtained in vacuum at the temperatures of 130, 155, 180, 205, 230, 255, 280, 305 and 330 K. No conformational transitions are recorded at this particular dihedral angle in a total of 27 ns runs (3 ns at each temperature). Jumps are recorded only for up to two of these Asp residue dihedrals even at very high temperatures (figure 3-9). The specific interactions between solvent and solute may be due to hydrogen bonds, as emphasized in previous studies [24, 32, 71], as well as other favorable interactions

between polar atoms. In fact, we have monitored solvent molecules near the catalytic site residues, and find that a close distance is maintained between the nearest oxygen atom of the solvent and the side-group heavy atoms; a collinear hydrogen bonding angle is not necessarily retained between the donor and acceptor groups.

In general, as the solvent molecules get more mobile with increasing temperature, they start to bump into the side chains, eventually causing them to make rotational jumps between their isomeric states. Moreover, the occurrence of these jumps increases at higher temperatures. Thus, inertial effects due to the type of solvent are also detrimental in the exact nature of the protein response. In figure 3-10, we display the time and temperature dependent torsional angle trajectories of two of the catalytic triad residues, Asp99 and Glu116, in the two solvents. Asp99, located on the basin of the catalytic site, makes conformational jumps below the transition temperature in water, whereas the same residue cannot do so below the transition temperature in glycerol.
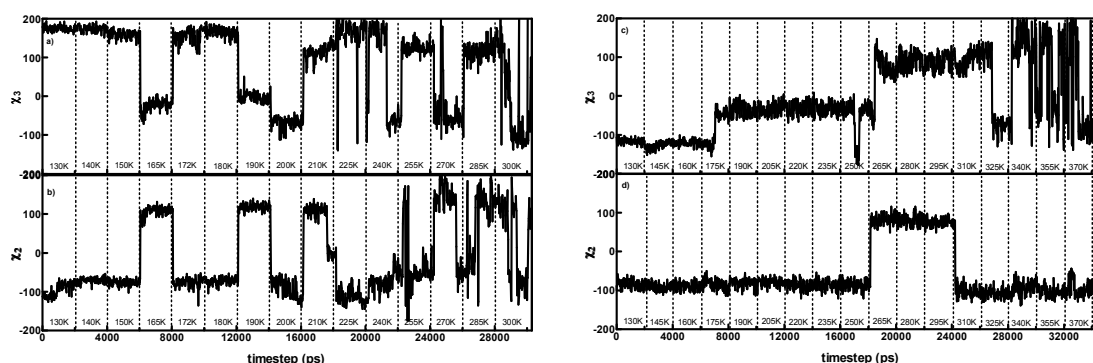


**Figure 3-10. Torsional angle trajectories of selected catalytic residues in different solvents over the whole temperature range studied. a) Glu 116 in water, b) Asp 99 in water, c) Glu 116 in glycerol, d) Asp 99 in glycerol. Note that the whole set of the MD simulations may be regarded as a heating procedure: The final structure from a simulation at a given temperature is the initial structure of the simulation at the next temperature. The trajectories resulting from the 0.5 ns equilibration periods (assumed to be transients between the structures equilibrated at the lower and higher temperatures) are not shown**

Since water has a small size, it may gain access to the catalytic site even at low temperatures, and occasionally gain enough energy to kick Asp99 so as to make it jump over the barrier separating the side-chain conformational states. Glycerol, on the other hand, cannot fit into the catalytic site due to its large size. It may gain access to the catalytic site and triggers Asp99 to make the jump only after the protein has enough flexibility to allow for the necessary volumetric fluctuations. This scenario is the same for Glu116 which is located at

the entrance of the catalytic site. Thus, larger solvents require a larger degree of coupling between solvent molecules and protein fluctuations in the vicinal layer. Note that the effect of solvent on the size or shape of the protein is a matter of discussion in the literature (see, e.g., [72] and references cited therein; Ansari et al. report that increasing glycerol concentration in solvents of glycerol-water mixtures has no perceptible effect on the size or shape of myoglobin [25] Here, the size of the protein remains essentially the same in both solvents, showing ca. 1% difference when averaged over all frames collected throughout the simulations which total to more than 40 ns in each solvent.

In fact, the transition temperature in glycerol is probably postponed until the system gains large enough fluctuations to accommodate the larger solvent molecules (the mean square fluctuation from figure 3-3 is 1.9 and 2.5 in water and glycerol at the point of their respective dynamical transitions). Once the vicinal layer forms with the establishment of the necessary communication between the solvent and solute, an even larger amount of fluctuations is introduced due to the processes that now become accessible on the potential energy surface leading to a marked increase in the flexibility.

The extent of coupling between librational motions and conformational jumps in chain molecules markedly affect the dynamics of the system. In general, the conformational entropy of the system is distributed between the backbone and side-chains. The force constant of the protein in glycerol is larger than that in water (eq.3.7 and figure 3-3) indicating a less flexible backbone once the system has gone through the dynamical transition. The protein in glycerol also exhibits a somewhat lower ability to make side-group torsional jumps. Furthermore, the stretch exponent, $\beta$, provides information about the distribution of time scales contributing to the relaxation of the protein in the solvent. In the protein-water system ($\beta = \sim 0.5$ in figure7), as opposed to the protein-glycerol system ($\beta = \sim 0.4$), a wider distribution of time scales, and therefore a larger variety of processes contributing to relaxation, is implied.

## 3.5 Nature of solvent−protein communication through radial distribution functions and favorable interactions.

The radial distribution function (RDF) gives the probability of finding an atom at a given distance *r* from another atom. It describes how, on average, the atoms in a system are packed [73].
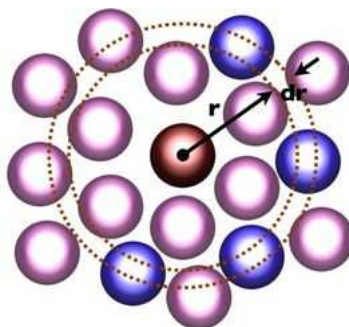


**Figure 3-11. Radial distribution functions using a spherical shell of thickness *dr*.**

To calculate the RDF from a simulation the neighbors around each atom or molecule in a shell at *r* + *dr* are counted (Figure 3-11). At regular intervals, a snapshot of the system is taken and the number of atoms found in each shell is counted. At the end of the simulation, the average number of atoms in each shell is calculated. This is then normalized by the volume of each shell and the average number density of atoms *ρ* in the system:

$$g(r) = n(r) / (\rho 4\pi r^2 \Delta r)$$

( 3-14 )

in which *g* (*r*) is the RDF, *n* (*r*) is the mean number of atoms in a shell of thickness *Δr* at distance *r*. The method need not be restricted to one atom type. All the atoms in the system can be treated in this way, leading to an improved determination of the RDF as an average over many atoms.

The RDF is usually plotted as a function of the interatomic separation *r*. A typical RDF plot (Figure 3-12) shows a number of important features. First, at short separations (small *r*) the RDF is zero. This indicates the effective size of the atoms, since they cannot approach any closer. Second, a number of obvious peaks appear, which indicate that the atoms pack around each other in coordination shells of neighbors. The occurrence of the peaks at long range indicates a high degree of ordering. Usually, at high temperature the peaks are broad,

indicating thermal motion, while at low temperature they are sharp. They are particularly sharp in crystalline materials, where atoms are strongly confined in their positions. At very long range the RDF is expected to approach to 1, implying that the bulk number density has been reached.
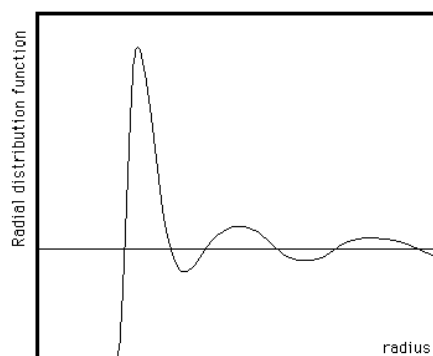


**Figure 3-12. A typical radial distribution function**

To investigate the influence of the solvent below and above the dynamical transition, selected RDFs and certain favorable interactions between solvent and protein atoms were also monitored. RDFs were computed for the distribution of oxygen atoms of the solvent molecules around the uncharged oxygen of the Asp 99 and Asp 126 residue (side chain). The RDF is averaged over the 1000 recorded snapshots of each trajectory. We also applied a third degree, five-point smoothing to the RDF data [47].

The RDFs for Asp 99 and Asp 126 are calculated by using VMD $g(r)$ GUI Plugin Version 1.0. Figure 3-13 displays the radial distribution between the solvent oxygen atoms and uncharged O of Asp 99 in water and glycerol, respectively. The $\Delta r$ is taken to be 0.1 Å and the RDFs are monitored up to 10 Å distance. Also, the data in the figure 3-13 is nudged upwards to get a clearer picture.
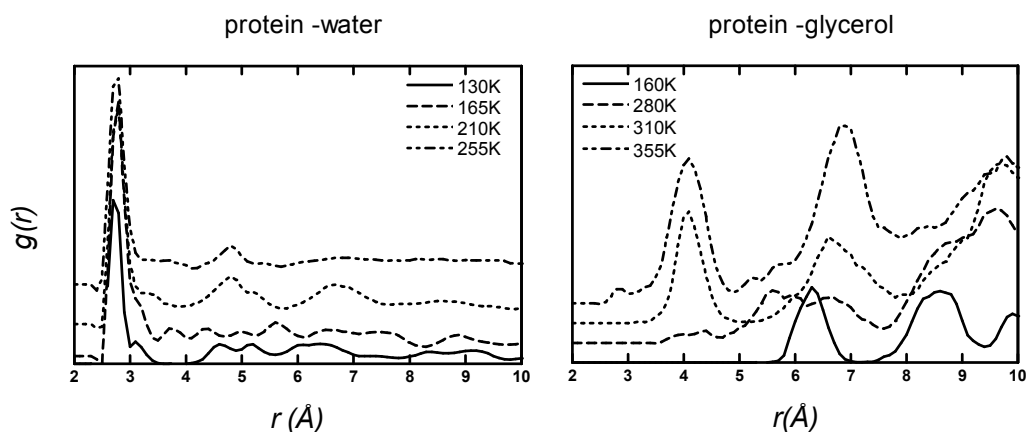
**Figure 3-13. Radial distribution functions of Asp 99 at selected temperatures in water and glycerol.**

The graphs reveal information about the orientation and mobility of the solvent below and above the transition. The catalytic site is a C-shaped cage (see figure 2-2), Asp 99 is located at its base. Since the size of water molecule is less than that of glycerol, more water molecules can go inside this cage whereas several glycerol molecules fill up some part of this cage.

In figure 3-13, we see more pronounced peaks at around 2.5, 4.5, 6.5 and 8.5 Å at 130 K for the RDFs of Asp 99 in water. The density of the latter peaks start to decrease as the temperature is raised and at 210 K these three peaks merge together to form two peaks, one at 4.5 Å and the other at 6.5 Å. Finally at 255 K the latter one vanishes and only the peak at 2.5 Å and 4.5 Å is left. On the other hand, in glycerol peaks at around 6.5 and 8.5 Å are seen at 160 K. These peaks start to melt and broaden at 280 K which is the temperature during transition in glycerol. This behavior is comparable to the results in water at 165 K the transition for the protein in water. Thus, during the transition the solvent molecules are rearranged and this arrangement takes place until the transition is completed. After transition we see peaks at 4, 6.5 and 9.5 Å. Note that the first peak appearing in glycerol moves to a closer distance when compared to the ones at lower temperatures. This is due to the fact that the size of the glycerol molecule is large and it only can enter to this cage when the protein becomes more flexible with increased temperature so as to accommodate the larger solvent. Concurrently, the solvent molecules gain mobility and at higher temperatures, with the help of energetic effects, they start to gain more energy and more movement ability around the protein surface. Finally at 355 K, the three peaks merge in to two appearing at 4 and 7 Å.

When, the solvent molecules inside the catalytic site cage are monitored we find that it always accommodates solvent molecules to some degree and as the system is heated up even the solvent molecules having a large size can fill up this cage. Thus, the protein rearranges the solvent during transition in such a way that it becomes functional. However, all of these results show the behavior of the solvent inside the catalytic site not the bulk solvent.

To understand the behavior of the bulk solvent, RDFs of another Asp (Asp 126) at the surface of protein is monitored at the same temperatures. Figure 3-14 shows the RDFs of Asp 126 in water and glycerol, respectively.
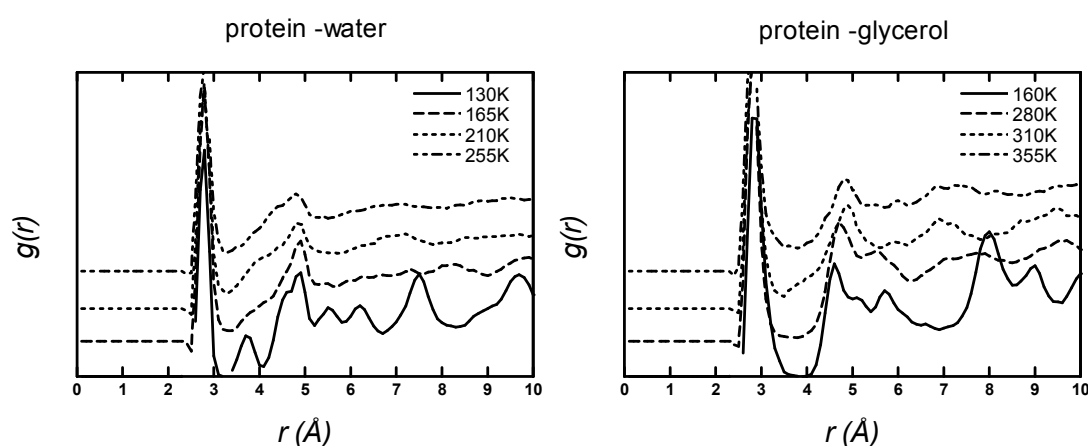


**Figure 3-14. Radial distribution functions of Asp 126 at selected temperatures in water and glycerol.**

Below transition, in water at 130 K many pronounced peaks are seen at around 2.5, 3.5, 4.5, 5.5, 6.5, 7.5 and 9.5 Å. Under the corresponding conditions in glycerol at 160 K, peaks are seen at around 3, 4.5, and 8 Å with a shoulder at 9 Å. These peaks start to merge together, and become less sharp as the temperature is raised due to thermal motion of solvent molecules. They become less ordered and more mobile as the system is heated up. Thus, the temperature dependent effect of the two solvents are the same on most of the protein where side chains are exposed, but they may have different effects on geometrically constrained regions.

To see the change in the mobility of the protein with temperature some snapshots were taken during the runs and were analyzed. The solvent molecules which are closer than 4 Å to the $C_\alpha$ atoms of the catalytic site residues were also monitored (figure 3-15). Two temperatures, one below and one above the transition were chosen for both of the solvents.

The solvent is more mobile above the transition and covers a much larger space around the protein surface. This result supports the observation of enhanced side-chain torsional angle jumps seen above the transition as outlined in the previous section. Since the solvent molecules gain more mobility thus flexibility as the temperature is raised, they start to kick and bump the side chains of the protein and the side chains make rotational jumps making the protein wiggle and jiggle (the slurry view in figure 3-15). As they sweep more space around the protein surface they trigger more side chains to display conformational transitions so that the solvent molecules cover even larger space. A continuous feedback is established between the solvent molecules, side-chain torsional jumps and backbone fluctuations.
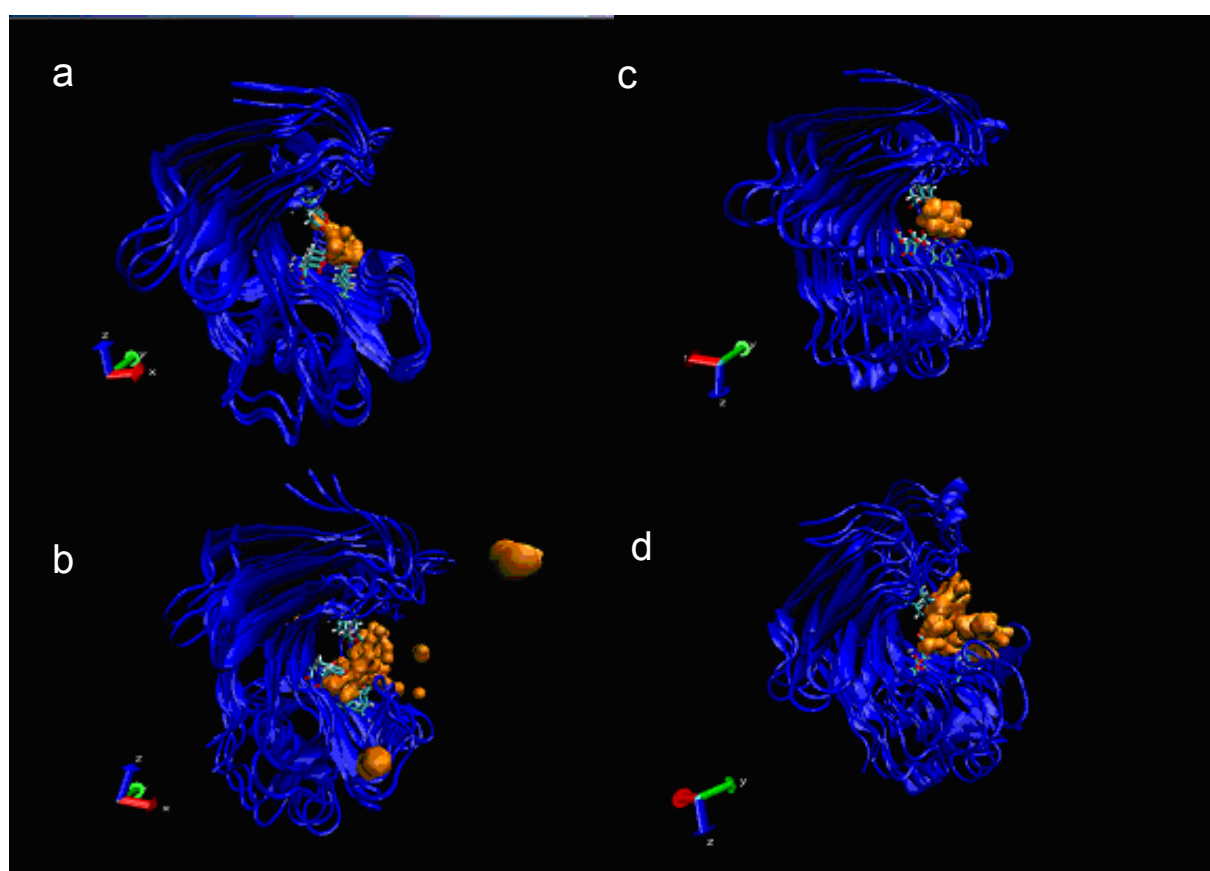


**Figure 3-15. 1h8v shown at 2, 1000, 2000 ps a) in water at 150 K. b) in water at 255 K. c) in glycerol at 250 K. d) in glycerol at 355 K. Solvent molecules are shown in surface representation (orange).**

To gain further insight into the molecular basis of the solvent effect, interactions between the uncharged O of the Asp 99 and solvent molecules are monitored at different temperatures. At each temperature, 10 snapshots are taken (100 ps intervals) and the solvent molecules around the uncharged O of Asp 99, which fall within the distance of the first peaks appearing

in figures 15 and 16 are selected. The distance between the hydrogen atoms of the solvent molecules and uncharged O of Asp 99 is monitored as well as the angle between the uncharged O of Asp 99 and H-O of the solvent molecule is monitored (Figure 3-16). Hydrogen bonding is defined when the distance is less than 2 Å, and the angle is greater than 150$^o$. All other solvents within the first coordination shell (less than 3.5 Å) of Asp 99 are labeled as favorable interactions.
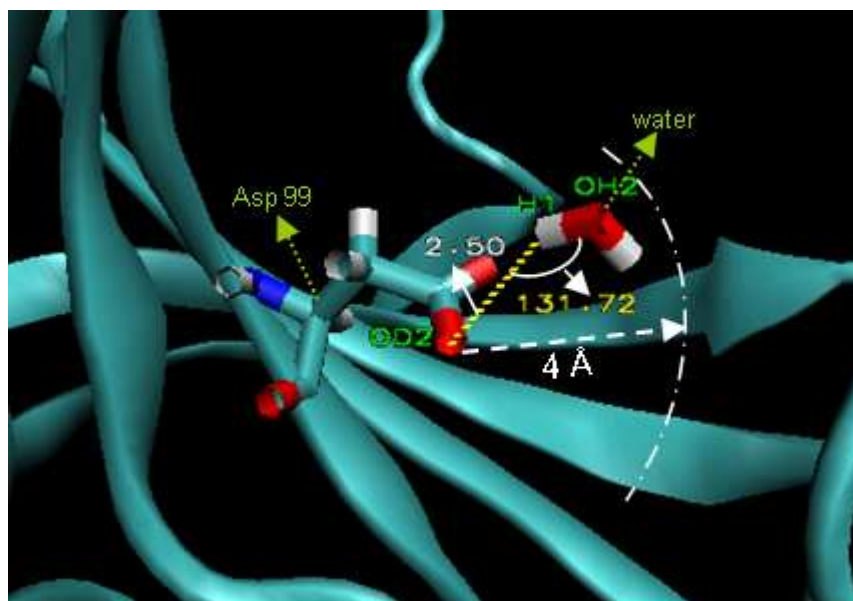


**Figure 3-16. The distance and angle measured to monitor the favorable interactions berween Asp 99 and nearby solvent molecules. The dotted circle shows the distance swept from the O of Asp 99; water is the one falling into this range. The distance between the O of Asp 99 and nearest H of water and the O (Asp 99)-H (water)-O(water) angle is measured.**

In figure 3-17, the distance vs. angle data for all these nearby water molecules are plotted. Each dot in the figure corresponds to one snapshot and different colors represent different water molecules or the same water molecule that flips around to interact with its other hydrogen atom. The hydrogen bonding interactions are boxed on the lower right part of the figures and the favorable interactions within the first coordination shell lie below the dotted line. At 130 K there are interactions which fall in hydrogen bonding so that these water molecules restrict the movement of the side chain. We find two such water molecules, one (red) is in the hydrogen bonded arrangement throughout the simulation whereas another (black) drifts from the hydrogen bonding region to larger distance and smaller angle region. During the transition at 165 K, different water molecules sample the region around the side chain more freely and make hydrogen bonds. Since the water molecules have changed they

find a new arrangement for making the interaction. Asp 99 now can accommodate three water molecules in the hydrogen bond arrangement simultaneously. As shown by the torsional angle trajectories in figure 3-10, the side chain takes a slightly different conformation, but does not make any jump during 165 K due to these strong hydrogen bonds.
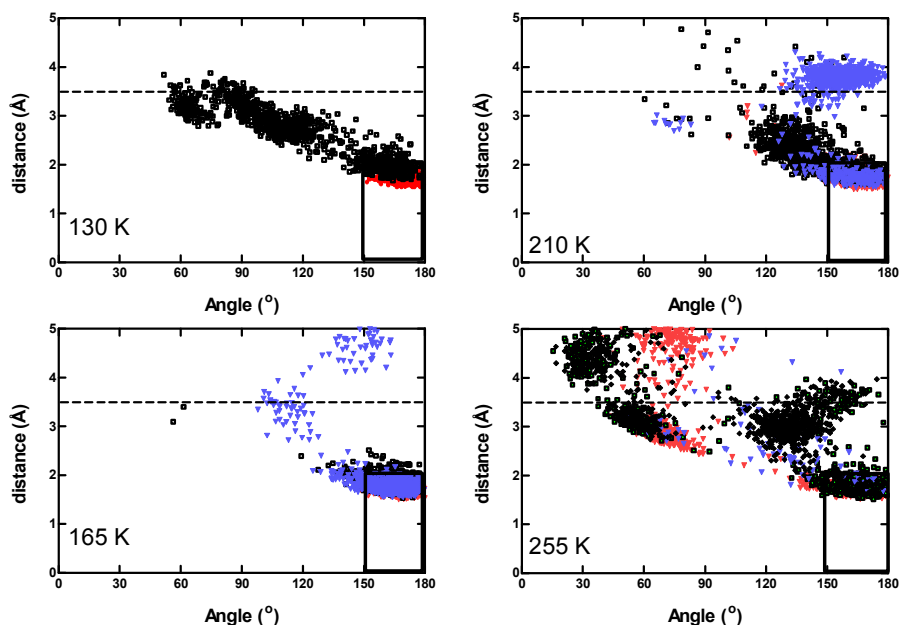


**Figure 3-17. The distance – angle trajectories of Asp 99 in water**

At higher temperatures, the mobility of the water molecules is even higher and the hydrogen bonding is frequently lost and sometimes regained but there are never more than three water molecules that get close to the Asp 99 residue. The results for the system in glycerol are shown in figure 20. As shown with the RDF for glycerol (figure 3-14), due to large size of this molecule, the catalytic cage is deprived of solvent at lower temperatures. However, as the temperature is raised, the cage gets more flexible and glycerol can go closer to Asp 99 and can fulfill the requirements of hydrogen bonding. Also note that at different temperatures, the data is for only one glycerol molecule, no other glycerol molecules could be found within the distance of 7.5 Å at 160 and 280 K and within 5 Å at 310 and 355 K. However, only the hydrogen of the glycerol which is close to Asp 99 uncharged O has changed by flipping. Also from figure 3-10 we see that Asp 99 in glycerol do not make jumps; this is probably due the glycerol molecule being very far away from the uncharged O

of the Asp 99 which can be seen from figure 3-18, underlining the fact that the solvent molecule triggers the torsional jumps.
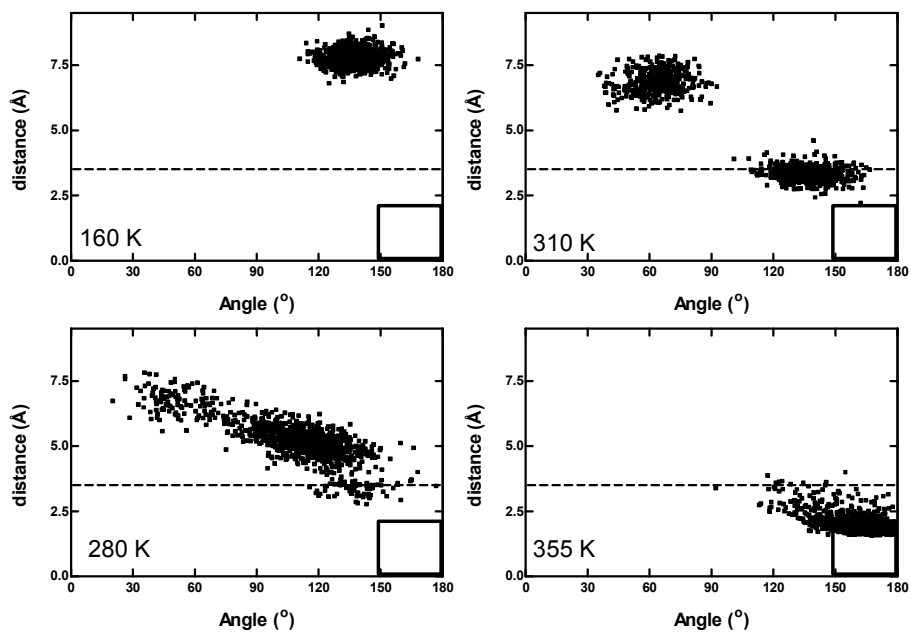
**Figure 3-18. The distance – angle trajectories of Asp 99 glycerol.**

# 4. Conclusions

In summary, the main findings of this work are the ones listed below. Each finding will be explained in detail in the following paragraphs.

- Different dynamical and thermodynamical methods are employed for depicting the dynamic transition of *1h8v* in water and glycerol.
    - Constant volume heat capacity
    - Mean positional fluctuations of the main-chain $C_\alpha$ atoms
    - Stretch exponential fit to the relaxation curves of the relaxation of the main-chain $C_\alpha$ atoms

- Below the transition the protein and solvent are decoupled from each other, completely acting on different time scales. Due to this fact, the relaxation behavior, the mean positional fluctuations and energy fluctuations of the protein in different solvents below the transition temperature are the same.

- At the onset of the transition, motions with intermediate time scales are introduced, with the coupling between the protein and a vicinal layer of the solvent. An additional frictional effect is introduced from the bulk solvent to this vicinal layer which in turn influences the protein dynamics.

- The average relaxation time of backbone fluctuations measured for the protein in both solvents and at all temperatures above the transition temperature is the same. This can be explained by the distribution of different time scales appearing during the coupling. The protein shows the same relaxation behavior independent of the environment implying that it is not the solvent but rather the protein itself which organizes its environment to achieve its functional state.

- A simple model based on Brownian dynamics with different stiffness and friction terms outlines all the observations made for the folded protein at a large variety of temperatures and in different solvents.

- A thermodynamical model is used to explain the shift in the transition temperatures in different solvents. It is due to the competition between the entropic and energetic factors that arise due to the interactions between the vicinal layer and the protein if they impinge similar types of interactions on the protein surface.

Using thermodynamical and viscoelastic arguments, we explain the nature of the protein dynamical transition and its dependence on different solvents by studying the equilibrium dynamics of the protein 1h8v in water and glycerol at a variety of temperatures. Mean positional fluctuations, a scalar which may be thought of as the fluctuations of the protein mass center, display a slope change at a critical temperature. When the critical temperature is reached, the vicinal water, which is treated as another unified mass coupled with the mass center of the protein, not only shifts the mass center, but also its fluctuations by weakening the effective spring constant (figure 3-3 and eq.3.7). We, metaphorically, state that, below the critical temperature, the mass center of the protein is decoupled from the solvent; i.e. vicinal solvent is indifferent from the bulk below the critical temperature. We depict this change in the states as the protein dynamical transition, sometimes referred to as the protein glass transition.

Relaxation time versus temperature graph also shows a state change at the critical temperature (figure 3-6). This is the relaxation time of the mass center's positional fluctuation, requiring damping or friction operating on the mass center of the protein (eq.3.5). Below the critical temperature, since the mass center of the protein independently communicates with bulk water, the latter impinges retardation on the protein referred to as $\zeta_p$. Above the critical temperature, a similar frictional effect, $\zeta_s$, from the bulk solvent also influences the vicinal layer. Furthermore, since an interaction is also switched on between the masses, symbolized by $k_{ps}$, concomitantly, $\zeta_{ps}$ is induced. This is a further change in the entropy due to the additional excluded volume effects that bulk solvent creates simultaneously on vicinal solvent and the protein.

The existence of $\zeta_{ps}$ can be qualitatively observed on the relaxation times. It introduces another time scale to the problem in the following way: $\zeta_p/k_p$ and $\zeta_s/k_s$ are the two scales of the protein and the vicinal layer, respectively, independent of each other. Their existence is due to the decoupled interactions of protein with the bulk and the vicinal layer with the bulk. Below the transition only the former exists, and the additional conformational motions are not induced, justified by the number of conformational transitions observed in the MD simulations. Right after the transition, another time scale is introduced mainly due to $\zeta_{ps}$, which modifies the relaxation time given approximately by $\tau_p \approx \zeta_p/k'_p$, if only the alteration of the $\zeta_p/k_p$ due to energetic interaction is considered. Yet, the concomitant entropic term is not there until $\zeta_{ps}$ is also considered and $\tau_p$ is further modified (eq. 3-9).

Thus, interaction energy and entropy is due to $k_{ps}$, which alter the mean positional fluctuation slope, and $\zeta_{ps}$ which affects the relaxation times by introducing another independent time scale. This effect is validated by the additional number of conformational transitions recorded from the MD simulations. Therefore, a unified Brownian dynamics with different stiffness and friction terms outlines all the observations made for the folded protein at a large variety of temperatures and in different solvents. The competition between the entropic and energetic factors that arise due to the interactions between the vicinal layer and the protein further explain the alteration in the value of the transition temperature in different solvents if they impinge similar types of interactions on the protein surface.

The details of the phenomena observed at the level of molecular detail are manifested in data reflecting the collective behavior of the system. The heat capacity of the system above the transition is less for the protein-glycerol than the protein-water system (figure 3). The smaller water molecules are less restricted to move around the side chains, compared to glycerol molecules, leading to a more entropic state for the protein-water system. This is also in agreement with our findings for residue fluctuations (figure 3-3), further supported with the torsional angle trajectories of figure 3-10. For the protein-water system, even before the onset of the transition, the side chains have a chance to make jumps due to high entropic contribution of water leading to small energy barriers, providing communication between different sub-state conformations. Apart from the monitoring of specific dihedrals, figure 3-4 shows that in general the jump rates increase with the increased mobility of the protein, more so in water than in glycerol.

Side-chain conformational transitions relieve excess energy that is stored in the system. If this relaxation pathway does not exist, the protein backbone torsional angles would be forced to make similar jumps. However, such jumps are known to require cooperativity between closely located dihedral angles along the backbone of chain molecules [74], so as to minimize the work done against the frictional environment during the displacement of the attached atoms [75]. In proteins, these conformational transitions are prohibitive in that they lead to unfolding of the chain. In fact, we have monitored the trajectories at temperatures close to unfolding, and find that these are earmarked by attempted, but short-lived, conformational jumps on the backbone (see, e.g. the 310 K $\psi$ trajectory in figure 4 of ref. [34]). The vicinal solvent layer closely interacts with the surface groups, and provides an alternate route for the system to spend the accumulated energy while maintaining the protein with enough flexibility to perform its function.

The average relaxation time of backbone fluctuations measured for the protein in both solvents and at all temperatures above the transition temperature is the same, in spite of the different distribution of conformational routes employed in different environments (as measured by the stretch exponent $\beta$, figure 3-5). Thus, the functional protein operates at a narrow time scale, and organizes its environment for achieving this operational state.

### *Future Work*

To understand and gain more insight into the origin of motions that provide the inserted time scales during transition, relaxation behavior of the selected side chains may be monitored. This study may also be repeated for other proteins to check if similar behavior is observed as parameters such as protein size or type (e.g. all alpha, all beta) are varied. Additionally, the dielectric properties and diffusion characteristics of the system may be studied to gain more information about the solvent effect on the protein. Furthermore, to validate the thermodynamical model which is proposed in this study, the same protein can be studied in different solvents.

Here, the proposed thermodynamical model is mainly based on the interactions between the solvent molecules and protein surface atoms, which is hydrogen bonding dominated for the solvents studied. This work can be repeated for (i) $D_2O$ instead of water in which no hydrogen bonding will occur, and (ii) deutrated glycerol in which the hydrogen bonding patterns are expected to be significantly altered. No rescaling of the dynamical transition temperature is expected in the former and an intermediated scaling is expected in the latter. Lastly, to inspect the protein in a dry environment, the simulations can be done with protein soaked in a bath of short peptides. By imitating an environment similar to the crowded cell medium, this approach might eliminate some of the caveats of performing simulations in vacuum, on in the crystal environment.

# 5. References

1.    Smolin, N., *A Simulation Analysis of Protein Hydration*. 2006, University of Dortmund: Dortmund.

2.    Karplus, M. and M. Andrew, *Molecular dynamics simulations of biomolecules*. Nature Structural Biology, 2002. **9**(9): p. 646-652.

3.    Colonna-Cesari, F., *Interdomain motion in liver alcohol dehydrogenase: structural and energetic analysis of the hinge bending mode*. Journal of Biological Chemistry, 1986. **261**: p. 15273-15280.

4.    Harvey, S.C., et al., *Phenylalanine transfer RNA: molecular dynamics simulation*. Science, 1984. **223**(1189-1191): p. 1189-1191.

5.    Case, D.A. and M. Karplus, *Dynamics of ligand binding to heme proteins*. Journal of Molecular Biology, 1979. **132**: p. 343-368.

6.    Brooks, B.R. and M. Karplus, *Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor*. Proc. Natl. Acad. Sci., 1983. **80**: p. 6571-6575.

7.    Irikura, K.K., et al., *Transition from B to Z DNA: contribution of internal fluctuations to the configurational entropy difference*. Science, 1985. **229**: p. 571-572.

8.    Ringe, D. and G.A. Petsko, *The 'glass transition' in protein dynamics: what it is, why it occurs, and how to exploit it*. Biophysical Chemistry, 2003. **105**: p. 667-680.

9.    Branden, C. and J. Tooze, *Introduction to Protein Structure*. Second Edition ed. 1991, Stockholm, Sweeden: Karolinska Institute.

10.   Parak, F., et al., *Dynamics of metmyoglobin crystals investigated bynuclear gamma resonance absorption*. Journal of Molecular Biology, 1981. **145**: p. 825-833.

11.   Alben, J.O., D. Beece, and S.F. Bowne, *Infrared spectroscopy of photodissociated carboxymyoglobin at low temperatures*. PNAS, 1982. **79**: p. 3744-3748.

12.   Doster, W., S. Cusack, and W. Petry, *Dynamical transition of myoglobin revealed by inelastic neutron scattering*. Nature, 1989. **337**: p. 754-756.

13.   Austin, R.H., et al., *Dynamics of ligand binding to myoglobin*. Biochemistry, 1975. **14**: p. 5355-5373.

14. Srajer, V., L. Reinisch, and P.M. Champion, *Investigation of laser-induced long-lived states of photolyzed MbCO.* Biochemistry, 1991. **30**: p. 4886-4895.

15. Nocek, J.M., J.S. Zhou, and S.D. Forest, *Theory and practice of electron transfer within protein-protein complexes: application to the multidomain binding of cytochrome c by cytochrome c peroxidase.* Chemisrty Review, 1996. **96**: p. 2459-2490.

16. Hartmann, H., et al., *Cnformational substates in a protein: structure and dynamics of metmyoglobin at 80 K.* PNAS, 1982. **79**: p. 4967-4971.

17. Jr, R.F.T., J.C. Dewan, and G.A. Petsko, *Effects of temperature on protein structure and dynamics : X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K.* Biochemistry, 1992. **31**: p. 2469-2481.

18. Loncharich, R.J. and B.R. Brooks, *Temperature dependence of dynamics of hydrated myoglobin.Comparision of force field calculations with neutron scattering data.* Journal of Molecular Biology, 1990. **215**: p. 439-455.

19. Steinbach, P.J. and B.R. Brooks, *Hydrated myoglobin's anharmonic fluctuations are not primarily due to dihedral transitions.* PNAS, 1996. **93**: p. 55-59.

20. Smith, J., K. Kuczera, and M. Karplus, *Dynamics of myoglobin: comparision of simulation results with neuron scattering spectra.* PNAS, 1990. **87**(1601-16053): p. 1601.

21. Demmel, F., et al., *Vibrational frequency shifts as a probe of hydrogen bonds: thermal expansion and glass transition of myoglobin in mixed solvents.* Europian Biophysical Journal, 1997. **26**: p. 327-335.

22. Diehl, M., et al., *Water coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering.* Biophysical Journal, 1997. **73**: p. 2726-2732.

23. Vitkup, D., et al., *Solvent Mobility and the protein 'glass' transition.* Nature Structural Biology, 2000. **7**(1): p. 34-38.

24. Dirama, T.E., G.A. Carri, and A.P. Sokolov, *Coupling between lyzosyme and glycerol dynamics: Microscopic Insights from molecular-dynamics simulations.* The Journal of Chemical Physics, 2005. **122**(24): p. 244910.

25. Ansari, A., et al., *The Role of Solvent Viscosity in the Dynamics of Protein Conformational Changes.* Science, 1992. **256**: p. 1796-1798.

26. Ansari, A., et al., *Protein states and proteinquakes.* PNAS, 1985. **82**: p. 5000-5004.

27. Bizzarri, A.R. and S. Cannistraro, *Molecular Dynamics of Water at the Protein-Solvent Interface.* Journal of Physical Chemistry B, 2002. **106**: p. 6617-6633.

28. Doster, W., et al., *Thermal Properties Of Water in Myoglobin Crystals and Solutions at Subzero Temperatures.* Biophysical Journal, 1986. **50**: p. 213-219.

29. Barron, L.D., L. Hecht, and G. Wilson, *The Lubricant of Life: A Proposal That Solvent Water Promotes Extremely Fast Conformational Fluctuations in Mobile Heteropolypeptide Structure.* Biochemistry, 1997. **36**(43): p. 13143-13147.

30. Pal, S.K., J. Peon, and A.H. Zewail, *Ultrafast surface hydration dynamics and expression of protein functionality: alpha-Chymotrypsin.* PNAS, 2002. **99**: p. 1763-1768.

31. Dastidar, S.G. and C. Mukhopadhyay, *Structure, dynamics, and energetics of water at the surface of a small globular protein: A molecular dynamics simulation* Physical Review E, 2003. **68**(2).

32. Tarek, M. and D.J. Tobias, *Role of Protein-Water Hydrogen Bond Dynamics in the Proten Dynamical Transition.* Physical Review Letters, 2002. **88**(12): p. 138101.

33. Baysal, C. and A.R. Atilgan, *Relaxation Kinetics and the Glassiness of Proteins: The Case of Bovine Pancreatic Trypsin Inhibitor.* Biophysical Journal, 2002. **83**: p. 699-705.

34. Baysal, C. and A.R. Atilgan, *Relaxation Kinetics and the Glassiness of Native Proteins: Coupling of Timescales.* Biophysical Journal, 2005. **88**: p. 1570-1576.

35. Tournier, A.L., J. Xu, and J.C. Smith, *Translational Hydration Water Dynamics Drives the Protein Glass Transition.* Biophysical Journal, 2003. **85**: p. 1871-1875.

36. Fenimore, P.W., et al., *Bulk solvent and hydration-shell fluctuations, similar to alpha and beta fluctuations in glasses,control protein motions and functions.* PNAS, 2004. **101**(40): p. 14408-14413.

37. Reat, V., et al., *Solvent dependence of dynamic transitions in protein solutions* PNAS, 2000. **97**(18): p. 9961-9966.

38. Leach, A.R., *Molecular Modelling Principles and Applications.* 1st ed. 1996: Pearson Education Limited.

39. Alder, B.J. and T.E. Wainwright, *Phase Transition fir a Hard-sphere System.* Journal of Chemical Physics, 1957. **27**: p. 1208-1209.

40. Rahman, A., *Correlations in the Motion of Atomsin Liquid Argon.* Physical Review 1964. **A136**: p. 405-411.

41. Elberly, D.H., *Game Physics.* 2004: Morgan Kaufman.

42. Swope, W.C., et al., *A computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters* Journal of Chemical Physics, 1982. **93**: p. 9042-9048.

43. Weiner, S.J., et al., *A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins* Journal of The American Chemical Soceity, 1984. **106**.

44. Brooks, B.R., et al., *CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations.* J. Comp. Chem., 1983. **4**: p. 187-217.

45. Jorgensen, W.L. and J. Tirado-Rives, *The OPLS [Optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin.* Journal of The American Chemical Soceity, 1988. **110**: p. 1657-1666.

46. Hermans, J., et al., *A consistent empirical potential for water-protein interactions.* Biopolymers, 1984. **23**: p. 1513-1518.

47. Allen, M.P. and D.J. Tildesly, *Computer simulation of liquids.* 1987, Oxford: Clarendron Press.

48. Frenkel, D. and B. Smith, *Understanding molecular simulation: from algorithms to applications.* 1996, San Diego: Academic Press.

49. Rapaport, D.C., *The art of molecular dynamics simulation.* 1995, Cambridge: Cambridge University Press.

50. Sadus, R.J., *Molecular simulation of fluids: theory,algorithms and object orientation.* 1999, Amsterdam: Elsevier Science B.V.

51. Berman, H.M., et al., *The Protein Data Bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

52. Sandgren, M., et al., *The X-Ray Crystal Structure of the Trichoderma reesei Family 12 Endoglucanase 3,Cel12A, at 1.9 A Resolution.* Journal of Molecular Biology, 2001. **308**: p. 295-310.

53. James C. Phillips, R.B., Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. , *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**: p. 1781-1802.

54. M. J. Frisch, G.W.T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M.

Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *Gaussian 03*. 2003, Gaussian, Inc., Pittsburgh PA.

55. Accelrys, I., *MATERIALS STUDIO*. 2002: San Diego.

56. Humphrey, W., Dalke, A. and Schulten, K., *VMD - Visual Molecular Dynamics*. Journal of Molecular Graphics, 1996. **14**(1): p. 33-38.

57. Pagnani, A., G. Parisi, and F. Ricci-Tersenghi, *Glassy transition in a disordered model for RNA secondary structure.* Physical Review Letters, 2000. **84**: p. 2026-2029.

58. Bicout, D.J. and G. Zaccai, *Protein Flexibility from the Dynamical Transition: A Force Constant Analysis.* Biophysical Journal, 2001. **80**: p. 1115 -1123.

59. Zaccai, G., *How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering.* Science, 2000. **288**: p. 1604-1607.

60. Tsai, A.M., D.A. Neumann, and L.N. Bell, *Molecular Dynamics of Solid-State Lysozyme as Affected by Glycerol and Water: ANeutron Scattering Study.* Bİophysical Journal, 2000. **79**: p. 2728-2732.

61. Gabel, G., et al., *The Influence of Solvent Composition on Global Dynamics of Human Butyrylcholinesterase Powders: A Neutron-Scattering Study.* Biophysical Journal, 2004. **86**: p. 3152-3165.

62. Stillinger, F.H., P.G. Debenedetti, and T.M. Truskett, *The Kauzmann paradox revisited.* Journal of Physical Chemistry, 2001. **105**: p. 11809-11816.

63. Kohlrausch, R., *Theorie des Elektrischen.Rückstandes in der Leidener Flasche.* Annual Physical Chemistry, 1974. **91**: p. 179-214.

64. Williams, G. and D.C. Watts *Non-symmetrical dielectric relaxation behavior arising from a simple emprical decay function.* Transactions of the Faraday Soceity, 1970. **66**: p. 80-85.

65. Qian, H., *Single-Particle Tracking: Brownian Dynamics of Viscoelastic Materials.* Biophysical Journal, 2000. **79**: p. 137-143.

66. Zaccai, G., *Proteins as nano-machines: dynamics–function relations studied by neutron scattering.* Journal OF Physics: Condensed Matter, 2003. **15**: p. S1673-S1682.

67. Cornicchi, E., et al., *Controlling the Protein Dynamical Transition with Sugar-Based Bioprotectant Matrices: A Neutron Scattering Study.* Biophysical Journal, 2006. **91**: p. 289–297.

68. Zhou, Y., D. Vitkup, and M. Karplus, *Native Proteins are Surface-molten Solids: Application of the Lindemann Criterion for the Solid versus Liquid State.* Journal of Molecular Biology, 1999. **285**: p. 1371-1375.

69. Dirama, T.E., et al., *Coupling between lyzosyme and trehalose dynamics: Microscopic insights from molecular-dynamics simulations.* The Journal of Chemical Physics, 2006. **124**: p. 034901.

70. Okan, O.B., *Glassiness and Coupling of Time Scales in Functional Proteins*, in *Engineering and Natural Sciences.* 2005, Sabanvi University: Istanbul.

71. Dirama, T.E., G.A. Carri, and A.P. Sokolov, *Role of hydrogen bonds in the fast dynamics of binary glasses of trehalose and glycerol: A molecular dynamics simulation study.* The Journal of Chemical Physics 2005. **122**(11): p. 114505.

72. Farnum, M. and C. Zukoski, *Effect of Glycerol on the Interactions and Solubility of Bovine Pancreatic Trypsin Inhibitor.* Biophysical Journal, 1999. **76** p. 2716–2726.

73. Hansen, J.P. and I.R. McDonald, *Theory of simple Liquids.* 1986, London Academic.

74. Baysal, C., B. Erman, and I. Bahar, *Contribution of Short-Range Intramolecular Interactions to Local Chain Dynamics.* Macromolecules, 1994. **27**: p. 3650-3657.

75. Baysal, C., et al., *Kinematics of Polymer Chains in Dense Medium. 4. Effect of Backbone Geometry and Application to Polybutadiene.* Macromolecules, 1996. **29**: p. 2980-2988.