

INFLUENCE OF ANTIFREEZE PROTEINS ON LOCAL WATER STRUCTURE DYNAMICS IN PRESENCE OF OSMOLYTES

ANAND NARAYANAN KRISHNAMOORTHY

*Institute for Computational Physics, Universität Stuttgart,
Allmandring 3, 70569 Stuttgart, Germany
anand@icp.uni-stuttgart.de*

(Received 11 March 2013; revised manuscript received 20 April 2013)

Abstract: Antifreeze proteins are synthesized by various organisms to enable their cells to survive subzero environment in the arctic and polar regions. These proteins produce a difference between the melting and freezing points termed as thermal hysteresis. The main objective of this study is to examine the dynamics of water molecules and hydrogen bonds at the protein-water interface of antifreeze protein using atomistic molecular dynamics simulations using GROMACS. For this work a prototype of AFP (antifreeze protein) from antarctic notothenioids (*ala-ala-thr* repeats) and a mutant which is not antifreeze active were generated using PRODRG server. The hydration dynamics results revealed that the retarded water dynamics in the AFP compared to its mutant was responsible for the antifreeze activity. Furthermore a considerable increase in antifreeze activity were observed for the AFP in presence of osmolytes. The mechanism of action were tested using preferential binding parameter derived from Kirkwood-Buff integrals.

Keywords: hydration dynamics, hydrogen bond lifetime, osmolytes, Kirkwood-Buff integral, preferential binding parameter

1. Introduction

1.1. Antifreeze proteins and glycoproteins

Antifreeze proteins (AFPs) and glycoproteins (AFGPs) are particular classes of proteins that suppress ice growth in organisms and thereby enable their survival in cold and subfreezing conditions [1]. These proteins provide protection for these organisms by depressing the freezing point of local water bodies, reacting in a non-colligative manner. The addition of AFPs at the interface between solid ice and liquid water inhibits the thermodynamically favored growth of the ice crystal. Ice growth is kinetically inhibited by AFPs covering the water-accessible surfaces of ice [2]. This difference between the freezing and melting temperature is referred to as thermal hysteresis. It is used as the characteristic measure of the antifreeze activity of an AFP [3]. Both glucoproteins and AFPs which are extracted from the

blood of polar fish, usually exhibit up to 2 degrees Celsius of thermal hysteresis. These types are called moderate AFPs whereas certain insects exhibit up to 5 degrees Celsius of thermal hysteresis, and these are classified as hyperactive antifreeze proteins. Thermal hysteresis can be measured experimentally using a custom made nanoliter osmometer.

AFP's have many potential applications [4, 5], including their use as additives to improve the quality and life of frozen food, as cryoprotective agents for organ and cell cryopreservation and also as chemical adjuvants to cancer cryosurgery and in development of transgenic plants and animals with increased tolerance to cold environment.

Early experimental models suggested that hydrogen bonding of AFP threonine residues to ice was the major driving force of AFP-ice association (De Vries 1974, Raymond and De Vries 1977) with AFP binding to ice via the threonine face. However, mutation experiments [6] revealed that hydrogen bonding was not the driving mechanism and that Van der Waals and hydrophobic interactions were suggested to play important roles in AFP binding. In this paper the actual mechanism of AFP binding is analyzed using atomistic the molecular dynamics simulation, and then compared with the experimental results (Havenith and Leitner 2003). Havenith experimentally studied the dynamics using terahertz spectroscopy and predicted that antifreeze activity was directly correlated with long range collective hydration dynamics and provided evidence of a new model for the proteins to survive under extreme conditions. The results suggest that the protein perturbs the aqueous solvent over long distances and retarded hydration dynamics in the large hydration shell does not favor freezing [7].

Antifreeze glycoproteins (AFGPs) are found in Antarctic notothenioids and in the northern cod. The remarkable diversity and distribution of AFPs suggest that different types have evolved recently in response to the sea level glaciation occurring 1–2 million years ago in the Northern hemisphere and 10–30 million years ago in Antarctica [8]. This independent development of similar adaptations is referred to as convergent evolution. There are two reasons why many types of AFPs are able to perform the same function despite their diversity. Figure 1 represents different classes of AFPs with their characteristics.

Although ice is uniformly composed of oxygen and hydrogen, it has many different surfaces exposed for binding. Different types of AFPs may interact with different surfaces. Although the five types of AFPs differ in their primary sequence of amino acids, when each folds into a functioning protein, they may share similarities in their three dimensional or tertiary structure that facilitates the same interactions with ice.

The preliminary proposed mechanism in which freezing point depression is achieved by an adsorption-inhibition mechanism [10] in which the proteins recognize and bind quasireversibly to an ice surface, thereby preventing growth of ice crystals. In this thesis we used an AFP-water model instead of creating an ice-water interface, since it is difficult to simulate ice using the molecular dynamics

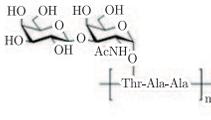
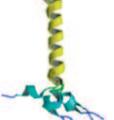
Characteristic	AFGP	Type I AFP	Type II AFP	Type III AFP	Type IV AFP
Mass (kDa)	2.6–33	3.3–4.5	11–24	6.5	12
Key properties	AAT repeat; disaccharide	Alanine-rich α -helix	Disulfide bonded	β -sandwich	Alanine-rich helical bundle
Representative structure					
Natural source	antartic notothenioids; northern cods	right-eyed flounders; sculpins	sea raven; smelt; herring	ocean pout; wolfish; eel pout	longhorn sculpin

Figure 1. Classes of AFPs and their characteristics [9]

simulation. With this system the actual mechanism is predicted and compared with the previous literature results. The PRODRG server is used [11] to create the topology and coordinate files.

1.2. Atomistic Molecular Dynamics simulations – GROMACS

Molecular Dynamics simulations (MD) is a technique for computing the equilibrium and transport properties of a classical many-body systems by solving Newton's equation of motions. This is an excellent approximation for a wide range of materials. The forces between the particles and potential energy are defined by molecular mechanics force fields. The actual measurement of properties was performed after equilibration. To measure an observable quantity in MD simulations, the observables must be expressed as a function of the position and momenta of the particles in the system. The MD simulation has proven useful in a large variety of molecular systems. Extended MD simulation was performed to characterize the motions and the conformational changes of the protein.

1.2.1. GROMACS overview

GROMACS is a versatile package to perform molecular dynamics, *i.e.* simulate the Newtonian equations of motion for systems with hundreds to millions of particles. It is primarily designed for biochemical molecules like proteins, lipids and nucleic acids that have a lot of complicated bonded interactions, however, since GROMACS is extremely fast at calculating nonbonded interactions that usually dominate simulations, many groups use it also for research on non-biological systems, *e.g.* polymers.

GROMACS is an acronym for Groningen Machine for Chemical Simulation. It was developed at the University of Groningen, the Netherlands, in the early 1990s. This open-source project is written in ANSI C, and contains about 100 utility and analysis programs which allow users to perform molecular simulations and energy minimization (EM) for biological molecules. It is one of the most commonly used molecular dynamics simulation packages. There are two levels of precision

that can be utilized in GROMACS – single and double. Some techniques, such as Normal Mode Analysis and more thorough energy minimization, require increased precision. In this work double precision is used.

1.2.2. Newton's laws of motions

Newton's laws of motion form the foundation of the branch of physics known as classical mechanics. It turns out, perhaps somewhat surprisingly, that Newton's laws can be applied, to a very good approximation, to study the motion of aggregates of molecules. The implication of this is that most atoms are heavy enough to treat their motion accurately within a classical framework. Obviously, the objects, being microscopic, the molecular motion is exactly described by the laws of quantum mechanics, and there are numerous instances in which the classical approximation breaks down. However, let us assume the approximate validity of classical mechanics at the microscopic level and introduce the basic laws of motion which will form the basis of our microscopic description of molecular systems. Mathematically it can be written as:

$$\vec{F} = \frac{d(m\vec{v})}{dt} = m \frac{d^2\vec{r}}{dt^2} \quad (1)$$

$$\vec{F}_{12} = -\vec{F}_{21} \quad (2)$$

where \vec{F} is the magnitude of force, m is the mass, \vec{v} is the velocity, \vec{r} is the position. The solution of these equations can, in principle, reproduce and predict the motion of any finite collection of massive objects. Molecular Dynamics simulations solve Newton's equations of motion for a system of N interacting atoms. Each force is the negative derivative of a potential function:

$$\vec{F}_i = -\frac{\partial u}{\partial \vec{r}_i} \quad (3)$$

The equations are solved for all atoms at times separated by a small time step Δt . There are many algorithms to integrate the equations of motion. In these simulations a leap-frog algorithm is used to solve the set of Newton's equation. The leap-frog algorithm uses positions r at time t and velocities v at time $(t - \delta t/2)$, it updates positions and velocities using forces $F(t)$ determined by the positions at time t using the relations:

$$\vec{v}(t + \delta t/2) = \vec{v}(t - \delta t/2) + \frac{\delta t}{m} F(t) \quad (4)$$

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t + \delta t/2) \quad (5)$$

The position update is given by:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \frac{\vec{F}(t)}{m} \delta t^2 + o(\delta t^4) \quad (6)$$

The algorithm is of third order in \vec{r} and is time-reversible.

1.2.3. Thermostat

Here, we simulate our system using an NVT ensemble, since our system is not isolated and exchange of energy is involved. In the NVT ensemble, the energy of endothermic and exothermic processes is exchanged using a thermostat.

Thermostats are designed to help a simulation sample from the correct ensemble (*i.e.* NVT) by modulating the temperature of the system. In simulations, the “instantaneous (kinetic) temperature” is usually computed from the kinetic energy of the system using the equipartition theorem [12]. In other words, the temperature is computed from the system’s total kinetic energy. The instantaneous value of the temperature is related to the kinetic energy via the particles momenta as follows:

$$\sum_i \frac{|P_i|^2}{2m_i} = \frac{k_b T}{2} (3N - N_c) \quad (7)$$

where N_c is the number of constraints. N represents the number of particles in the system. The average temperature T_i is identical to the macroscopic temperature.

Unfortunately, the microcanonical ensemble does not correspond to the conditions under which most experiments are carried out. If one is interested in the behavior of a system at a specific temperature, an NVT-ensemble simulation using a thermostat is required. Another reason to simulate using a thermostat is to avoid steady energy drifts caused by the accumulation of numerical errors during MD simulations. An obvious way to alter the temperature of the system is velocity scaling. If the temperature at time t is $T(t)$ and the velocities are multiplied by factor λ , then the associated temperature change can be calculated as [13]:

$$\Delta(T) = \frac{1}{2} \sum_i 2m_i \frac{(\lambda \vec{v}_i)^2}{N_d k_B} - \frac{1}{2} \sum_i 2m_i \frac{(\vec{v}_i)^2}{N_d k_B} \quad (8)$$

$$\Delta T = (\lambda^2 - 1)T(t) \quad (9)$$

$$\lambda = \sqrt{\frac{T_o}{T(t)}} \quad (10)$$

The simplest way to control the temperature is thus to multiply the velocities at each time step by factor λ where $T(t)$ is the current temperature as calculated from the kinetic energy and T_o is the desired temperature. One problem with this approach is that it does not allow fluctuations in temperature which are present in the canonical ensemble.

A weaker formulation of this approach is the Berendsen thermostat [14]. The system is coupled to an external heat bath with fixed temperature T_o to maintain the temperature. The velocities are scaled at each step, such that the rate of change of temperature is proportional to the difference in temperature:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_o - T(t)) \quad (11)$$

where τ is the coupling parameter which determines how tightly the bath and the system are coupled together. This method gives an exponential decay of the

system towards the desired temperature. The change in temperature between successive time steps is:

$$\Delta T = \frac{dt}{\tau} (T_o - T(t)) \quad (12)$$

Thus, the scaling factor for the velocities is:

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left(\frac{T_o}{T(t-\delta t/2)} - 1 \right) \quad (13)$$

The $T(t-\delta t/2)$ is due to the fact that the so called leap-frog algorithm is used for the time integration. In practice, τ is used as an empirical parameter to adjust the strength of the coupling. Its value has to be chosen with care. In the limit τ tending to infinity, the Berendsen thermostat is inactive and the run is sampling a microcanonical ensemble. The temperature fluctuations will grow until they reach the appropriate value of a microcanonical ensemble. However, they will never reach the appropriate value for a canonical ensemble. On the other hand, too small values of τ will cause unrealistically low temperature fluctuations. If τ is chosen the same as the time step δt , the Berendsen thermostat is nothing else than simple velocity scaling. The values of $\tau \approx 0.1$ ps are typically used in MD simulations of condensed-phase systems. The ensemble generated when using the Berendsen thermostat is not a canonical ensemble [13].

A velocity rescaled Berendsen thermostat is used in our simulation. It is very much similar to the Berendsen thermostat with an additional stochastic term which ensures a correct kinetic energy distribution by modifying it accordingly:

$$dK = K_o - K \frac{dt}{\tau} + 2 \sqrt{\frac{K K_o}{N_f}} \frac{dW}{\sqrt{\tau}} \quad (14)$$

where K is the kinetic energy, N_f the number of degrees of freedom and dW a Wiener process [15]. There are no additional parameters, except for a random seed. This thermostat produces a correct canonical ensemble and still has the advantage of the Berendsen thermostat, first order decay of temperature deviations and no oscillations.

1.2.4. Interaction functions and force fields

The force field refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles (typically molecules and atoms). Depending on the force field, the parameters may differ for each kind of atom as in an all-atom force field, or it may be more coarse-grained. Due to the wide range of goals in a molecular simulation, there are a variety of force fields with different strengths and weaknesses, tailored for one or more particular applications. The accuracy of an all-atom force field is tempting to the user, but the computational cost of large biomolecular simulations where intramolecular motion is less important than intermolecular motion makes coarse-graining an appealing concept.

There is no single force field that is superior to all others in all situations and it falls to the researcher to determine the appropriate force field for a given

system. Notable force fields include OPLS, GROMOS, CHARMM and AMBER [16]. Here we use gromos43a1 forcefield, which is a united atom forcefield. In addition to the coordinates and velocities, the force on every atom has to be calculated to determine the dynamics of the system. This is done by means of the following expression:

$$\vec{F} = -\nabla V \tag{15}$$

V is a semi-empirical potential function for which several expressions and sets of parameters, altogether called the force field, have been proposed.

A typical molecular force field has the following form:

$$V = V_{\text{bonded}} + V_{\text{non-bonded}} + V_{\text{special}} \tag{16}$$

The potential functions can be subdivided into three parts:

1. Non-bonded: Lennard-Jones or Buckingham, and Coulomb or modified Coulomb. The non-bonded interactions are computed on the basis of a neighbor list (a list of non-bonded atoms within a certain radius) in which exclusions have been already removed:

$$V_{\text{non-bonded}} = V_{\text{L-J}} + V_{\text{Coulomb}} \tag{17}$$

$$V_{\text{L-J}} = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] = \epsilon \left[\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^6 \right] \tag{18}$$

where ϵ is the depth of the potential well, σ is the finite distance at which the inter-particle potential is zero, r is the distance between the particles and r_m is the distance at which the potential reaches its minimum. At r_m the potential function has the value $-\epsilon$. The distances are related as $r_m = 2^{1/6}\sigma$. The Coulombic interaction between two charged particles is given by:

$$V_{\text{Coulomb}} = \frac{K_e |q_i \cdot q_j|}{\epsilon_r r_{ij}}, \tag{19}$$

where q_i and q_j are the charges of particles i and j , $K_e = \frac{1}{4\pi\epsilon_0} = 138.93$. In GROMACS the relative dielectric constant ϵ_r can be set in the input of `grompp`.

2. Bonded: covalent bond-stretching, angle-bending, improper dihedrals, and proper dihedrals [16]. These are computed on the basis of fixed lists:

$$V_{\text{bonded}} = V_{\text{bonds}} + V_{\text{angles}} + V_{\text{improper}} + V_{\text{proper}} \tag{20}$$

Bonded interactions are based on a fixed list of atoms. They are not exclusively pair interactions, but include 3- and 4-body interactions as well. There are bond stretching (2-body), bond angle (3-body), and dihedral angle (4-body) interactions [16].

3. Restraints: position restraints, angle restraints, distance restraints, orientation restraints and dihedral restraints, all based on fixed lists.

1.2.5. Structural details of alanine and threonine

Proteins are large biomolecules consisting of one or multiple chains of amino acids. Peptide bonds are formed from two amino acid bases. Here (*ala-ala-thr*)₁₅

protein dynamics is performed. The protein consists of repeated units of alanine-alanine-threonine subunits. This usually forms the binding chain in almost all antifreeze proteins. The structural formula for alanine is given in Figure 2 and threonine in Figure 3.

The α -carbon atom of alanine is bound with a methyl group ($-\text{CH}_3$) Figure 2, making it one of the simplest α -amino acids with respect to the molecular structure and also resulting in alanine being classified as an aliphatic amino acid. The methyl group of alanine is non-reactive and is thus almost never directly involved in the protein function. Alanine is nonpolar and it is one of the hydrophobic amino acid bases due to the presence of aliphatic methyl side chain. However, it can play a role in substrate recognition or specificity, particularly in interactions with other non-reactive atoms such as carbon.

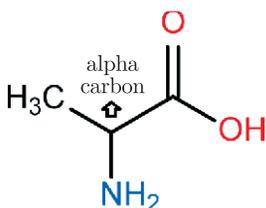


Figure 2. Structural formula of alanine

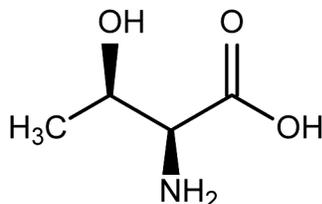


Figure 3. Structural formula of threonine

Threonine is an α amino acid which has a hydroxyl group attached thereto. It also bears a chiral side chain. Threonine is classified as polar amino acid and it is hydrophilic. The threonine residue is susceptible to numerous posttranslational modifications. The hydroxyl side-chain can undergo O-linked glycosylation. Hence in AFGPs the glycosidic linkages are attached to the threonine side chain. The hydroxyl group is fairly reactive, being able to form hydrogen bonds with a variety of polar substrates. Being a fairly indifferent amino acid, Threonine can reside both within the interior of a protein, or on the protein surface. It is fairly neutral with regard to mutations, though generally it substitutes other polar or small amino acids.

1.3. Introduction of cosolutes to protein-water system

It has been studied that antifreeze activity can be enhanced by the presence of polyhydroxy and polycarboxylate compounds [17]. Also the presence of a citrate buffer in an AFP-water solution has shown increased antifreeze activity [18].

Therefore, in this thesis we study the effect of urea and the hydroxyectoine presence in an AFP-water system.

Protein in a solution exists in a conformational equilibrium with an ensemble of unfolded states, with the folded states being favored at ambient conditions. An equilibrium between folded and unfolded ensembles can be perturbed by changing the composition by adding cosolvents to the solution [19]. The addition of a cosolvent tends to stabilize or destabilize the conformation of the protein. Urea and hydroxyectoine (Figures 4–5) were chosen as cosolvents with water to study the effect of stabilization of proteins.

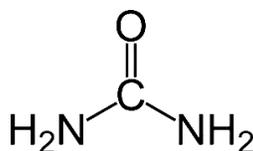


Figure 4. Structural formula for urea

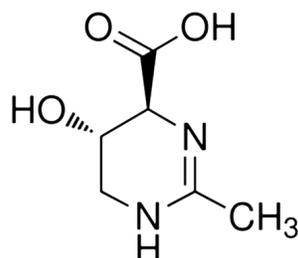


Figure 5. Structural formula for hydroxyectoine

The terms kosmotrope (order-maker) and chaotrope (disorder-maker) originally denoted solutes that stabilize, or destabilize proteins. These terms refer to the apparently correlating property of increasing, or decreasing the structuring of water (a recent review of structure-making and structure-breaking ions has been presented [20]). The classification of a solvent into a chaotrope or a kosmotrope is difficult sometimes, since certain compounds behave differently at varying concentrations. For example, a solute may not always act in the same way at different concentrations or in the presence of macromolecules or gels [21]. Also some solutes with less well-defined properties like urea are sometimes classified as kosmotropes [21] and sometimes as chaotropes [22] depending on concentration. In this paper we choose a 6M concentration of both hydroxyectoine and urea. With such concentrations urea acts as a chaotrope and hydroxyectoine as a kosmotrope [23].

An alternative term used for kosmotrope is 'compensatory solute' as they have been found to compensate for the deleterious effects of high salt contents (which destroy the natural hydrogen bonded network of water) in osmotically stressed cells. A review of the literature shows that hydroxyectoines are small

molecular weight compatible solutes which are zwitterionic and tend to bind more to water molecules compared to urea [24]. In this paper we study and predict these underlying mechanism using atomistic molecular dynamics simulations.

The biological importance of these solutes lies in stabilization of proteins. Urea generally tends to destabilize a protein structure by a direct mechanism under high concentration. The direct mechanism suggests that urea interacts directly with the protein backbone, via hydrogen bonds, and other electrostatic interactions predominantly interact with charged and polar side chains [25]. Whereas the hydroxyectoine molecule tends to stabilize protein using an indirect mechanism in which the solute does not interact directly with the macromolecule, but tends to preferentially exclude from the protein surface [26, 27]. In this paper we study and predict these underlying mechanisms using atomistic molecular dynamics simulations.

2. Computational details

All the simulations were carried out with GROMACS, version 4.5.5, in double precision. Typically the simulations sets are split into two parts, one being simulations of AFPs and its mutant with water as the solvent and the other set – with osmolytes as cosolutes and water as the solvent. The water model chosen in this work is the SPC/E water model and all the charges are parameterized with the gromos43a1 force-field. A brief description of the protein simulation using GROMACS is given in the present section.

2.1. Protein topology generation using PRODRG server

PRODRG is server software used for generation of topologies for small molecules [11]. The manual generation of topologies is time consuming and error prone, and hence, this server is proved useful in creating prototypes of [(*ala-ala-thr*)₁₅] AFP. Figure 6 depicts the AFP modeled by the PRODRG server.

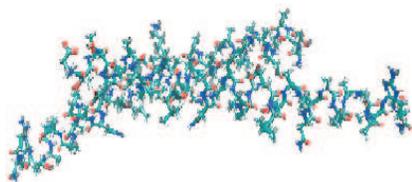


Figure 6. AFP generated from PRODRG

The PRODRG server generates molecular force field geometries about small molecules. The server can take up the initial configuration or any pdb file as input to generate these topologies. This server deletes all the hydrogen atom information and ignores all the bond order information. These steps use chemical pattern matching and hydrogen generation by proper orbital hybridization of carbon atoms. Currently all hydrogens are generated by 17 rules [28]. Most of the steps in PRODRG use chemical pattern matching. Hydrogen is generated by

identifying the corresponding hybridization of C, N, O atoms. The server is flexible and user configurable [11], and can generate topologies compatible with GROMACS force-fields (gromos43a1). PRODRG supports only atom types limited to C, H, N, O, P, S, F, Cl, Br and I. An atom cannot be connected to more than four other atoms and this version is limited to a maximum of 600 atoms. Also, ignoring the hydrogens and bond types may lead to unexpected results.

All the topology and coordinate files were generated using the PRODRG server. The initial configuration of (*ala-ala-thr*) was sketched using the server itself, as was its mutant (*ala-ala-ala*). The charges for the protein chain obtained from the topology files were edited according to the charges compatible with gromos43a1 force-fields. The resulting topology file had some errors in the partial charges, those were edited during the course of simulation to make it compatible with the gromos43a1 force fields. Also the osmolytes such as hydroxyectoine and urea were generated using the PRODRG server. These names and types of residues were edited during the course of simulation.

2.2. Protein-water simulation using GROMACS

The suitable topology and coordinate files were opted out using the gromos43a1 force-field. In the next few steps, the box size was defined and the protein was introduced and it was solvated in the box. The GROMACS tool `editconf` is very useful to change the format of coordinate files, to rotate and translate coordinate files and to define the box size:

```
editconf -f afp.gro -o newbox.gro -c -d 1.0 -bt cubic
```

The above command centers the chain protein in the box (`-c`), and places it at least 1 nm from the box edge (`-d 1.0`). The box type is defined as a cube (`-bt cubic`). The distance to the edge of the box is an important parameter. Since periodic boundary conditions are used, the minimum image convention must be satisfied. That is, a protein should never see its periodic image, otherwise the forces calculated will be enormous. Specifying a solute-box distance of 1 nm will mean that there are at least 2 nm between any two periodic images of a protein. This distance will be sufficient for just about any cutoff scheme commonly used in simulations. Now that protein is introduced in a cubic box, water molecules must be added for solvation of the protein. This is performed by the `genbox` operation:

```
genbox -cp newbox.gro -cs spc216.gro -o solv.gro -p topol.top
```

The configuration of the protein (`-cp`) is contained in the output of the previous `editconf` step, and the configuration of the solvent (`-cs`) is part of the standard GROMACS installation. In this case the space water model is used which is compatible with the gromos43a1 force fields which are applied to these simulations, `genbox` keeps a track of how many water molecules it has added, which it then writes to the topology to reflect the changes that have been made.

2.2.1. Energy minimization

The system is now solvated and it is electrically neutral. Before the dynamics is performed the system must not have any steric clashes due to

inappropriate geometry. Hence, it is necessary to minimize the energy of the system. The system is energy minimized using the steepest descent algorithm for 50 000 steps with the maximum force limit less than 100 kJ/mol/nm.

```
grompp -f minim.mdp -c solv.gro -p topol.top -o em.tpr
```

The command `grompp` was used to process the topology and parameter files provided as an input to the system. This command is used to generate an atomic level input file `.tpr` which is used to perform the sampling run and this input file contains the details of all the kinds of force applied and the interaction details as well:

```
mdrun -v -deffnm em
```

The `mdrun` command was used to run the sampling simulation run. The energy minimization is performed for 200 ps. The two important factor that needs to be verified are the force and potential energy obtained from this run. The force value must be less than 1000, since the restraints are applied.

Energy minimization ensures a reasonable starting structure, in terms of geometry and solvent orientation. The dynamics starts with the equilibration process. The restraints are applied to prevent the system collapse. Equilibration is performed to make the real dynamics of the system. It is used to sustain the system to the required temperature in this case.

In the canonical ensemble, the number of molecules N , volume V and temperature T are conserved. It is also sometimes called constant temperature molecular dynamics (CTMD). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat.

The simulations were performed with a modified Berendsen thermostat as described earlier. The simulations were performed at 300 K:

```
grompp -f nvt.mdp -c em.gro -p topol.top -o nvt.tpr
mdrun -deffnm nvt
```

As the system was now well equilibrated for the required temperature and physical conditions, the next step was to run the production simulation run to create the suitable trajectory files for protein-water dynamics analysis:

```
grompp -f md.mdp -c nvt.gro -t nvt.cpt -p topol.top -o md.tpr
mdrun -deffnm md
```

Similarly to the equilibration, the production runs were performed for 20 ns. The initial steps were omitted to stabilize the system and then the rest of it was used for analysis.

2.3. Performed simulations

2.3.1. AFP in water

Several sets of simulations were performed. The first set of simulations was performed for the $(ala-ala-thr)_{15}$ chain derived from the PRODRG server solvated in water. This chain was solvated using SPC/E water model in a cubic box.

Then, the system was energy minimized using the steepest descent algorithm to obtain a relaxed system. The box was equilibrated in an NVT ensemble using a V-rescale Berendsen thermostat. Then, the system was allowed to develop the trajectory files using the production run for 20 ns. The cubic box ($5 \times 5 \times 5$) nm³ of simulation comprised of 3964 molecules. For mutation studies, the *(ala-ala-ala)*₁₅ chain was simulated under similar conditions as the *(ala-ala-thr)*₁₅ chain. The box was generated with ($5 \times 5 \times 5$) nm³ 4080 water molecules. The solvent density for all the simulations was identical.

The next set of simulations was performed to analyze the concentration effect on hydration dynamics. Two chains of *(ala-ala-thr)*₁₅ were solvated in SPC/E water and the simulations were done for 20 ns in an NVT ensemble with the same set of parameters as above. For analyzing the temperature effects the AFP chain was simulated in water at 275 K in an NVT ensemble using a Berendsen thermostat for 20 ns. All these sets of simulations were done with the 2 fs time step and the electrostatic interactions were carried out by the Particle Mesh Ewald (PME) method, in-built in GROMACS-4.5.5.

2.3.2. AFP in water with cosolutes (hydroxyectoine and urea)

In this case the generated AFP chain was solvated in a 6M urea water solution and for the second set the system was solvated in a 6M hydroxyectoine water solution, respectively. This system was also energy minimized using the steepest descent algorithm, and then equilibrated using the NVT ensemble with a V-rescale Berendsen thermostat. The production run was performed for 20 ns at the 2 fs timestep at 300 K as well. The simulation set for AFP in hydroxyectoine-water was performed in a cubic box ($4.69 \times 4.69 \times 4.69$) nm³ in size and with 168 hydroxyectoine and 1581 water molecules. The AFP in urea-water simulation was performed at ($4.69 \times 4.69 \times 4.69$) nm³ cubic box with 237 urea molecules and 3205 water molecules. Figure 7 shows the simulation box of an AFP in aqueous urea, where red denotes water molecules and the protein is in the center (gray) surrounded by an urea molecule (green).

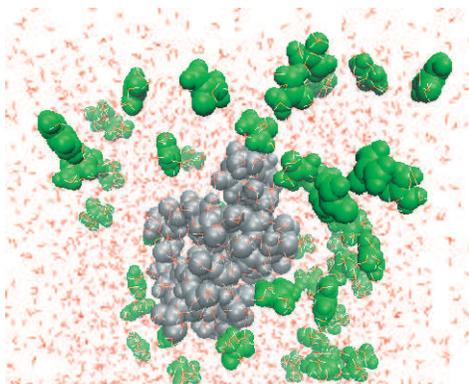


Figure 7. Snapshot of AFP chain in aqueous urea solution from GROMACS

Similar MD simulations were performed for the mutation chain as well for a comparative study. To analyze the preferential binding and exclusion studies, both alanine and threonine bases were simulated separately in a 6M urea water, and then 6M hydroxyectoine water solution at an NVT ensemble at 300K using a Berendsen thermostat and the 2fs timestep. The simulation to analyze the electrostatic effect was done by removing all the partial charges from the protein, and then simulating with the same parametric set.

2.4. Results analysis

2.4.1. Radial distribution function

In statistical mechanics, the radial distribution function (RDF or pair correlation function) $g(r)$ in a system of particles (atoms, molecules, colloids, *etc.*), describes how density varies as a function of distance from a reference particle. If a given particle is taken to be at the origin O , and if $\rho = N/V$ is the average number density of particles, then the local time-averaged density at a distance r from O is $\rho g(r)$. This simplified definition holds for a homogeneous and isotropic system. A more general case will be considered below. In simplest terms it is a measure of the probability of finding a particle at a distance of r away from a given reference particle, relative to that for an ideal gas. The general algorithm involves determining how many particles are within a distance of r and $r + dr$ away from a particle. The RDF is usually determined by calculating the distance between all particle pairs and binding them into a histogram. The histogram is then normalized with respect to an ideal gas, where particle histograms are completely uncorrelated. For three dimensions, this normalization is the number density of the system multiplied by the volume of the spherical shell, which mathematically can be expressed as $g(r)_I = 4\pi r^2 \rho dr$, where ρ is the number density. The radial distribution function (RDF) $g_{AB}(r)$ between particles of type A and B is defined in the following way:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{local}} = \frac{1}{\langle \rho_B \rangle_{local}} \frac{1}{N_A} \sum_{N_A} \sum_{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2} \quad (21)$$

with $\langle \rho_B(r) \rangle$ the particle density of type B at a distance r around particles A , and ρ_{Blocal} the particle density of type B averaged over all spheres around particles A with radius r_{max} . Usually the value of r_{max} is half of the box length. The averaging is also performed in time. In practice the analysis program `g rdf` divides the system into spherical slices and makes a histogram, which can be interpreted as a δ function.

2.4.2. Hydrogen bond dynamics

The program `g hbond` analyses the hydrogen bonds (*H-bonds*) between all possible donors D and acceptors A . To determine if an *H-bond* exists, a geometrical criterion is used:

$$r \leq r_{HB} = 0.35 \text{ nm} \quad (22)$$

$$\alpha \leq \alpha_{HB} = 30 \text{ deg} \quad (23)$$

The value of $r_{HB} = 0.35$ nm corresponds to the first minimum of the RDF of SPC water. The `g hbond` program analyses all hydrogen bonds existing between two groups of atoms (which must be either identical or non-overlapping) or in specified donor-hydrogen-acceptor triplets, in the following ways:

- Donor-Acceptor distance (r) distribution of all *H-bonds*;
- Hydrogen-Donor-Acceptor angle (α) distribution of all *H-bonds*;
- The total number of *H-bonds* in each time frame;
- The lifetime of the *H-bonds* is calculated from the average over all autocorrelation functions of the existence functions (either 0 or 1) of all *H-bonds*:

$$C(\tau) = \langle s_i(t)s_i(t+\tau) \rangle \quad (24)$$

with $s_i(t) = (0,1)$ for *H-bond* i at time t . The integral of $C(\tau)$ gives a rough estimate of the average *H-bond* lifetime τ_{HB} :

$$\tau_{HB} = \int C(\tau) d\tau \quad (25)$$

Both the integral and the complete autocorrelation function $C(\tau)$ will be the output, so that more sophisticated analysis (*e.g.* using multi exponential fits) can be used to get better estimates for τ_{HB} .

The value for the lifetime allows an estimate of relative strength for the corresponding hydrogen bond [29, 30]. The lifetime is related to the activation energy by:

$$K_f = \frac{1}{\tau_{HB}} = \frac{K_B T}{h} \exp\left(\frac{-\Delta G}{K_B T}\right) \quad (26)$$

where τ_{HB} is the hydrogen bond lifetime, K_B is the Boltzmann constant, h is the Planck constant, and ΔG gives the activation energy. This shows that a larger lifetime corresponds to larger free activation energy, hence, strengthening the hydrogen bond network.

Another important quantity that can be measured is the number of present hydrogen bonds n_{HB} . Also the hydrogen bond density (ρ_{HB}) [31] can be obtained by dividing n_{HB} by total solvent accessible surface area (σ_t). This counts the average number of hydrogen bonds per unit surface area.

2.4.3. Preferential binding parameter

Proteins are seldom solvated by pure water. Other solvent components, such as buffer salts and stabilizers, are ubiquitous in the laboratory and in formulations of therapeutic proteins. Similarly, intracellular solutions are crowded with many types of proteins, metabolites, nucleic acids, osmolytes, and other molecules. The presence of these other components, hereafter called cosolvents, generally alters protein equilibria and reaction kinetics by perturbing the chemical potential of the protein system. Cosolvents perturb the chemical potential of the protein system by associating either more strongly or more weakly with the protein than water. This phenomenon, called preferential binding, is of great interest because it governs the physical and chemical properties of proteins. In the literature a number of different

definitions of the preferential binding parameter have been employed. They can be connected by thermodynamic relations for ternary mixtures. The preferential binding coefficient is a way in which binding can be defined thermodynamically. It is also particularly useful when binding is weak. The preferential binding coefficient is a measure of the excess number of cosolvent molecules in the domain of the protein per protein molecule [32].

The preferential binding parameter can also be derived from the Kirkwood-Buff theory of solutions. Several authors have reported results in this direction [33, 34]. The relation is given as (1 – represents water, 2 – represents protein and 3 – represents cosolvent):

$$\Gamma_{23} = c_3(G_{23} - G_{13}) \quad (27)$$

where G_{13} , G_{23} are the Kirkwood-Buff integrals defined as:

$$G_{\alpha\beta} = \int_0^\infty (g_{\alpha\beta} - 1) 4\pi r^2 dr \quad (28)$$

where $g_{\alpha\beta}$ is the radial distribution function between species α and β , and r is the distance between the centers of molecules α and β .

As a rule, $\Gamma < 0$ the cosolvent is preferentially excluded from the protein, for example cosolvents such as glycerol, sucrose, *etc.*, can stabilize protein structure at high concentrations preserve its enzymatic activity, and for $\Gamma > 0$ the cosolvent is preferentially binded to the protein (such as urea), which can cause protein denaturation.

2.4.4. Transfer energy

The transfer model is a model used to study the influence of cosolvents on the protein stability. In this model, the folding/unfolding of the protein and the process of the transfer of folded and unfolded states from water to aqueous urea solution are tied together in a thermodynamic cycle. According to the transfer model, the free energy of protein unfolding in a 6M urea solution (ΔG_{6M}^{Urea}) is related to the free energy of unfolding in water (ΔG^{Water}) through:

$$(\Delta G_{6M}^{Urea}) = (\Delta G^{Water}) + \sum_i n_i \alpha_i \delta g_i^{tr} \quad (29)$$

Here, α_i is the fractional change in the solvent-accessible surface area of group i upon unfolding, δg_i^{tr} is the experimentally measured free energy of transfer of group i from water to a 6M urea solution, and n_i is the number of groups of type i present in the protein [35].

The transfer free energy or the chemical potential can be calculated using the preferential binding coefficient using the following relation:

$$\mu_p^{tr} = -RT \Gamma_{23} \quad (30)$$

This equation provides a simple and convenient link between preferential binding coefficients and free energies. Here, R is the gas constant, T is the temperature and Γ_{23} is the preferential binding parameter and μ_p^{tr} is the transfer free energy.

Experimental and theoretical studies show that [32]:

- Γ may be positive or negative, indicating that interactions of the protein and cosolvent are favorable or unfavorable, respectively;
- Γ is proportional to cosolvent molality at low concentration of cosolvent (often as high as molar concentration of cosolvent > 1 M and higher);
- Γ is roughly proportional to the protein-solvent interfacial area.

2.4.5. Dipole autocorrelation function

The dipole autocorrelation function for water molecules is defined as:

$$C_{\mu}(t) = \langle \mu(0) \cdot \mu(t) \rangle \quad (31)$$

where $\mu_i(t)$ is the dipole vector of water molecule i , and the angular brackets denotes averaging over initial times. The correlation function $C_{\mu}(t)$ is fitted using a stretched exponentially decaying function: $C_{\mu}(t) = A \exp(-t/\tau)^{\beta}$, where A represents the normalization constant and τ represents the average time constant (relaxation time). The decrease in the number of free orienting dipoles may be related to some of the water molecules binding to the protein backbone. The stretched exponential potential is dependent on the β value. The evaluation of τ allows a quantitative determination of water dynamics.

The dipole autocorrelation function $\langle \vec{\mu}_i(t) \cdot \vec{\mu}_i(t') \rangle$ describes the correlation between the electric dipole moment vector of a molecule at one time and the electric dipole vector of the molecule at a different time. The dielectric relaxation spectrum of a material is related to this correlation function. The dipole moment vector of a molecule is related to the orientation of the molecule. The mean autocorrelation time is given by:

$$\langle \tau \rangle = \int_{t_0}^{\infty} dt \exp(-t/\tau)^{\beta} = \tau/\beta \Gamma(1/\beta) \quad (32)$$

The direct connection between the entropy and the relaxation dynamics has been studied previously [36] in which it has been shown that the relation:

$$\tau = \tau_0 \exp(-S/k_B) \quad (33)$$

with the entropy S is valid for several measurable relaxation times.

2.4.6. Solvent accessible surface area

It is defined that the surface area of the biomolecule is accessible to a solvent. It is typically calculated using the rolling ball algorithm developed by Shrake and Rupley in 1973 [37]. This algorithm uses a sphere (of solvent) of a particular radius to “probe” the surface of the molecule (Figure 8).

`g sas` command in GROMACS computes the hydrophobic, hydrophilic and total solvent accessible surface area. The solvent accessible surface area is calculated by the sum of spheres centered at the atoms of the studied molecule, such that a spherical solvent molecule can be placed at a closest distance and in agreement to the van der Waals interaction by following the constraints that other atoms are not penetrated [38].

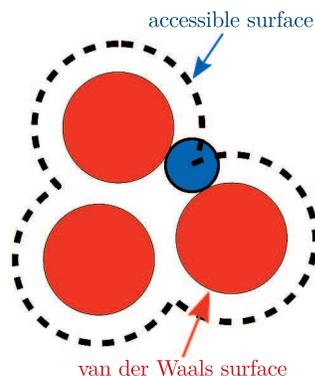


Figure 8. Illustration of the solvent accessible surface in comparison to the van der Waals surface; the van der Waals surface as given by the atomic radii is shown in red; the accessible surface is drawn with dashed lines and is created by tracing the center of the probe sphere (in blue) as it rolls along the van der Waals surface; the probe radius depicted here is of smaller scale than the typical 1.4 Å [37]

2.4.7. Radius of gyration

The calculation of radius of gyration gives a measure of compactness of a protein structure. In GROMACS it is calculated using `g gyrate`:

$$R_g = \sqrt{\frac{\sum_i |r_i|^2 m_i}{\sum_i m_i}} \quad (34)$$

where m_i is the mass of atom i and r_i the position of atom i with respect to the center of mass of the molecule. It is especially useful to characterize polymer solutions and proteins.

3. Results and discussion

The antifreeze protein is obtained from prototype antarctic notothenioid toothfish, *dissostichus maswoni* consisting of $(\text{ala-ala-thr})_{15}$ chains. Mutagenesis experiments have questioned the previous assumption of hydrogen bonding as the dominant mechanism and experimental results from terahertz spectroscopy show that the antifreeze activity is directly correlated with long range collective hydration dynamics. The results of Havenith's studies show that the AFP perturbs the aqueous solvent over long distances and the retarded water dynamics in this hydration region does not favor freezing. The phenomenon has been studied with molecular simulations and the obtained MD results are completely compatible with the experimental studies.

3.1. Brief account on experimental studies

Havenith's experimental studies were based on antifreeze glycoprotein – water dynamics. The thermal hysteresis was measured using a Clifton nanoliter cryoscope. In Figure 9 it is clearly seen that the value of thermal hysteresis increases with an increase in the AFGP concentration in water. The protein yielded

a maximum value of 1.7 degrees Celsius of thermal hysteresis and the hysteresis was considerably reduced in the presence of a borate buffer. It is also seen that at a certain AFGP concentration the thermal hysteresis curve reaches a saturation point which means that further addition of protein has no effect on the freeze resistance.

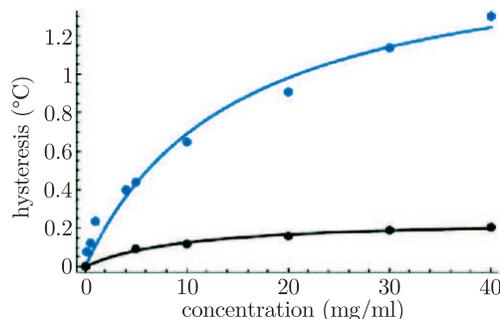


Figure 9. Hysteresis of AFGP (blue) and AFGP in 0.3M borate buffer (black) as a function of AFGP concentration in water [7]

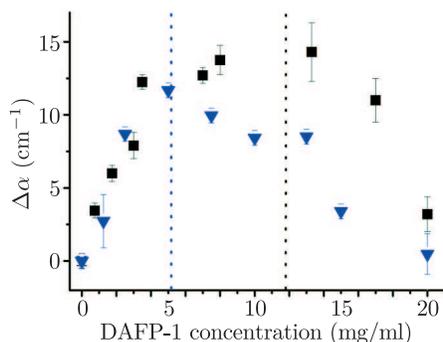


Figure 10. Concentration-dependent terahertz absorption of DAFP-1. Measurements were carried out at 293.15 K (20°C) (black data points) and 278.15 K (blue data points). DAFP-1 shows a concentration-dependent excess of terahertz absorbance that shifts to lower protein concentration at the lower temperature (blue *vs.* black dashed vertical lines) [18]

Leitner and Havenith studied the long range dynamics of a DAFP-water system using terahertz spectroscopy. Figure 10 shows the difference, $\Delta\alpha$, in the terahertz absorbance of DAFP-1 dissolved in water. The largest $\Delta\alpha$ value was observed to be 15/cm. When decreasing the temperature a shift of c_{\max} from ca. 12 to ca. 5 mg/ml was observed. This shift indicates an increase in the dynamic hydration shell at lower temperatures. The observation of long distance modification of dynamics at low temperatures shows optimal protein-water coupling. This retarded water dynamics in a large hydration shell does not favor freezing.

The abovementioned experimental results were directly correlated with the computational results obtained by simulating DAFP in SPC/E water at 300 K.

Figure 11 shows a plot for the hydrogen bond lifetime correlation function of water molecules around different DAFP protein planes. This shows a clear trend of difference in water dynamics in the vicinity of binding and non-binding planes of DAFP. The function $c(t)$ is decaying over longer times for the binding plane in comparison to the non binding planes. This heterogeneous distribution of hydrogen bond lifetimes between water molecules and protein for different planes of DAFP is correlated with antifreeze activity [39]. In general, a trend of retardation of hydrogen bond dynamics is witnessed.

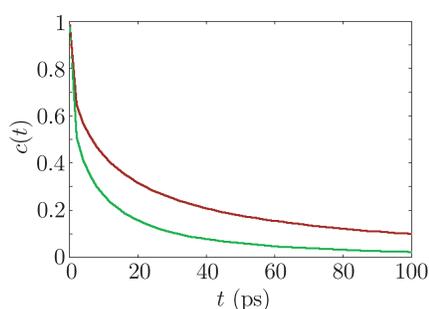


Figure 11. Hydrogen bond autocorrelation function for water molecules around DAFP binding plane (red) and DAFP non-binding plane (green)

Here, we study similar hydration dynamics for an antifreeze prototype of $(ala-ala-thr)_{15}$ as mentioned above. We correlate the obtained computational results with the experimental results. This prototype chain forms the major part of the ice-binding chain for various classes of AFPs [40].

3.2. Hydrogen bond dynamics of AFP around water

Antifreeze glycoprotein is constructed by addition of disaccharide galactosamine to the hydroxyl oxygen of threonine residue through a glycosidic linkage. The structure of AFGPs is very flexible compared to native AFPs. Figure 12 gives a schematic view of an AFGP generated from PRODRG.

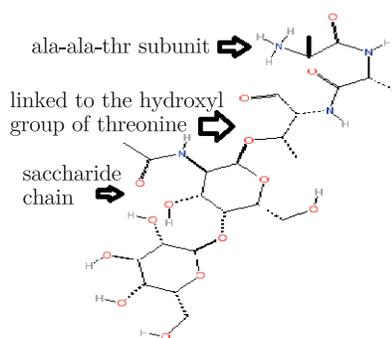


Figure 12. Structural representation of AFGP (Antifreeze Glycoprotein) from PRODRG server

In Figure 13 hydrogen the bond correlation function for water molecules around the AFP chain, $c(t)$, is plotted. The figure depicts the function up to 50 ps. It is compared with the hydrogen bonds within the water molecules as a bulk (blue). All the simulations are carried out at 300 K.

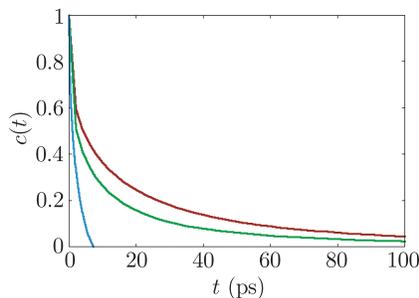


Figure 13. Hydrogen bond correlation function for water molecules around AFP chain (red) and its mutant (green) and bulk water (blue) at 300 K

The observed slow rearrangement times for hydrogen bonds of water in the bulk are consistent with the expectations for water molecules in the hydration layer around a protein. This retarded water dynamics in the large hydration shell does not favor freezing [7, 18]. The figure shows that correlation function $c(t)$ decays over longer times for water molecules near the (*ala-ala-thr*) plane compared to the hydrophobic (*ala-ala-ala*) chain [41] and water molecules in a bulk. This effect is due to the presence of threonine residue in AFP. Replacing threonine by alanine reduces the retardation effect of water around proteins.

Table 1 shows the rate constant [42, 43] obtained from the autocorrelation function which shows a lower value for the (*ala-ala-thr*) chain compared to the (*ala-ala-ala*) chain, further supporting our study. These values validate the above discussions. These higher lifetimes of AFP compared to mutant can be correlated to higher activation energy, thus showing a stronger hydrogen bonding network for water around AFPs compared to its mutant.

Table 1. Hydrogen bond autocorrelation function for bonds between water and protein

Peptide Chain	Rate Constant K_f (ps^{-1})	Lifetime t (ps)
AFGP in water	0.284	3.526
AFP in water	0.565	1.770
Alanine chain in water	0.816	1.226

The measures of hydrogen bonding at the protein-water interface help in quantifying the complexity and heterogeneity of the interactions between water and the antifreeze protein, and reveal regions of the protein-water interface important for the antifreeze activity.

Figure 14 shows a plot for the hydrogen bond correlation function for water molecules around AFGP and AFP. Compared to native AFPs, AFGPs exhibit high

flexibility and do not have any well defined structure. The effect of glycosidic linkages in the protein backbone also increases the antifreeze activity. This effect is shown by an increase in the hydrogen bond lifetime of water molecules around an AFGP in comparison with an AFP (Figure 14). The slower retardation dynamics of the AFGP may be due to the presence of a disaccharide chain. From this the presence of hydrophilic, polar groups that enhance antifreeze activity can be concluded.

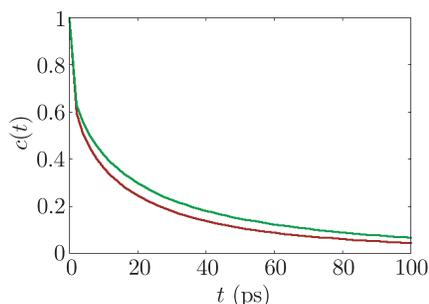


Figure 14. Hydrogen bond correlation function for AFP chain (green) and AFGP (glycosidic linkage to threonine base) (red) in water at 300 K

There is a significant increase in the antifreeze activity, when the temperature is lowered (Figure 15). This result is in complete agreement with the experimental results [18]. The plot of the hydrogen bond autocorrelation function for water molecules around the (*ala-ala-thr*) chain at 275 K decays over longer times compared to the AFP chain at 300 K.

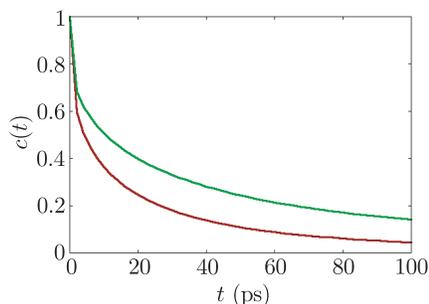


Figure 15. Hydrogen bond correlation function for AFP chain in water at 275 K (red) and 300 K (green)

Figure 16 shows the plot of hydrogen bond correlation function of water molecules around (*ala-ala-thr*)₁₅ and 2 chains of (*ala-ala-thr*)₁₅, to study the concentration effect on antifreeze activity. Even a concentration dependent analysis (Figure 16) of the AFP chain was performed using the simulation of two (*ala-ala-thr*) in the SPC/E water model. The AFP chain shows increased antifreeze activity with an increase in concentration [7].

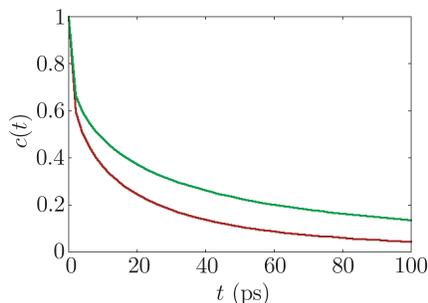


Figure 16. Hydrogen bond correlation function for $(ala-ala-thr)_{15}$ (red) and 2 chains of $(ala-ala-thr)_{15}$ (green) chain in water at 300K

Hence, the results of both the mutation and AFP studies signify that the presence of the hydrophilic threonine group in the chain is responsible for the increase in the antifreeze activity. The water molecules tend to bind towards the protein plane that contains threonine like a hydrophilic base. The experimental results show that the position of threonine in the protein backbone and the presence of the OH group increases the size of the hydration shell, thereby increasing the thermal hysteresis [44]. Also the site-specific mutation has a clear nonlocal effect on the active ice-binding plane, therefore, this heterogeneous hydrogen bond dynamics facilitates the suppression of the freezing point [45].

It can be noticed in Figure 17 that the number of hydrogen bonds formed between the $(ala-ala-thr)$ chain and water is much greater than in the mutant $(ala-ala-ala)$ chain. It is correlated with a simultaneous increase in the solvent accessible surface area of the AFP chain as shown in Figure 18. Hence, this effect can be partly related to the simple fact that the presence of the hydroxyl group in the bulkier and hydrophilic threonine group forms an increased binding property compared to its mutant. This effect completely supports the previous results obtained for the correlation function.

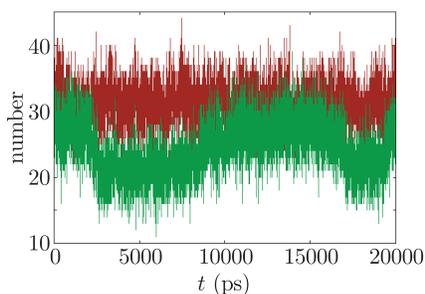


Figure 17. Number of hydrogen bonds as function of time for AFP (red) and its mutant (green) at 300K

Table 2 shows that the net average value of the number of hydrogen bonds (n_{HB}) between water and the AFP chain is larger than the mutant chain. Comparing these values from the Table, it is clear that the number of hydrogen

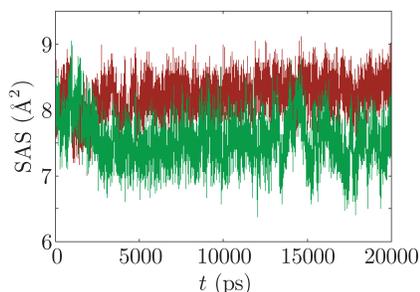


Figure 18. Solvent accessible surface area of AFP (red) and its mutant (green) at 300K

bonds increases with the increasing total solvent accessible (σ_t) surface area resulting in an increase in the hydrogen bond density (ρ_{HB} in Table 2). The AFP chain in comparison with the mutant chain can be correlated to the presence of polar threonine residue in the AFP chain.

Table 2. Hydrogen bonding parameters for water molecules around protein chain

Peptide chain	n_{HB}	σ_t (nm ²)	ρ_{HB} hydrogen bond density (nm ⁻²)
(<i>ala-ala-ala</i>)	24.474	12.099	2.022
(<i>ala-ala-thr</i>)	31.209	12.725	2.4525

3.3. Influence of cosolutes in enhancing antifreeze activity

There are several studies which state that AFPs show an increased activity on the addition of osmolytes. The molecular mechanism for this increase is still a matter of discussion. Hence, the AFP chains were investigated in aqueous urea and aqueous hydroxyectoine.

Previous studies of AFPs in the presence of cosolutes (using terahertz spectroscopy) have predicted a corresponding increase in the range of the retarded hydration shell and also found a corresponding increase in thermal hysteresis from 1.2°C to 6.8°C in the presence of a 0.5M citrate buffer at the same AFP concentration [18]. The mechanism of action is predicted to be the same as for AFP in water, namely, the presence of enhancers increases the antifreeze activity by increasing the long range dynamic retarded hydration shell which does not favor freezing. Also, once the maximum concentration of enhancers is reached, the increase in the antifreeze activity is saturated, showing maximum TH activity [43, 46].

Kosmotropic and chaotropic co-solvents influence the stability and biochemical equilibrium in aqueous solutions of proteins, acting indirectly through the structure and dynamics of the hydration water that surrounds the protein molecules [20]. This thesis depicts the study of the AFP chain in the presence of aqueous hydroxyectoine and aqueous urea. This optimal choice was made to investigate the effect of protein dynamics in the presence of kosmotrope (hydrox-

yectoine) and chaotrope (urea) as cosolutes with water as solvent. Urea behaves both as a kosmotrope and chaotrope under varying concentration [21], hence for a comparative study, the simulations were performed at a 6M concentration of hydroxyectoine and urea in water, respectively.

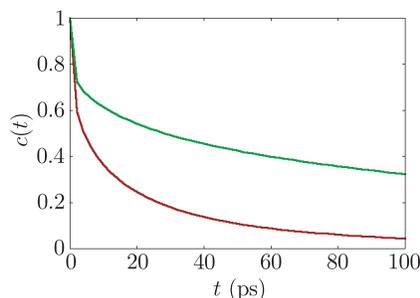


Figure 19. Hydrogen bond correlation function for AFP chain in presence of hydroxyectoine (red) and AFP chain (green) in water at 300 K

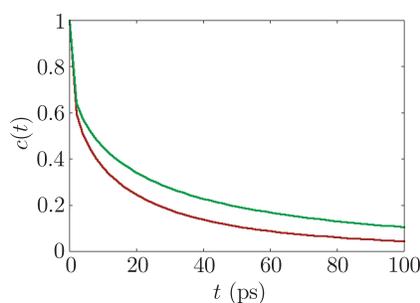


Figure 20. Hydrogen bond correlation function for AFP chain in urea (red) and AFP chain (green) in water at 300 K

Figure 19 represents the hydrogen bond lifetime correlation function for water molecules around the AFP chain in the presence of aqueous hydroxyectoine. The Figure clearly shows increased retardation dynamics for AFP in the presence of 6M hydroxyectoine as compared to the action of the AFP chain in water. Thus, the presence of hydroxyectoine influences the hydrogen bond lifetimes and increases the time for hydrogen bond rearrangement on the AFP chain. This effect of slowing down of hydrogen bond rearrangement dynamics of water near the protein chain in aqueous hydroxyectoine compared to protein in water is responsible for the increase in the high antifreeze activity of AFP [18].

Table 3. Hydrogen bond autocorrelation function for bonds between water and protein

Peptide chain	Rate constant K_f (ps^{-1})	Lifetime (ps)
AFP in water	0.565	1.770
AFP in aqueous urea	0.290	3.449
AFP in aqueous hydroxyectoine	0.126	7.922

Figure 20 shows the hydrogen bond lifetime correlation function for water molecules around AFP chain in the presence of aqueous urea. Similar effects are observed: enhanced retardation dynamics for the AFP chain in aqueous urea than in water. This slow rearrangement of hydrogen bonds in water molecules around the AFP chain is responsible for an increase in the antifreeze activity. Figure 21 shows that the hydrogen bond rearrangement times for water molecules around the protein chain were slower for the AFP chain in the presence of aqueous hydroxyectoine compared to aqueous urea, showing that the aqueous hydroxyectoine buffer results in producing greater antifreeze activity compared to aqueous urea. Hence, it is verified that the AFP chain behaves differently in the presence of different enhancers. Certain enhancers tend to effect larger thermal hysteresis compared to others, which also depends on concentration [7, 18]. For example the citrate buffer tends to increase thermal hysteresis to three folds at 0.5M concentration [18].

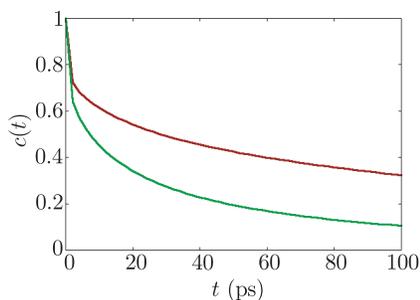


Figure 21. Hydrogen bond correlation function for AFP chain with urea (red) and AFP chain with hydroxyectoine (green) in water at 300 K

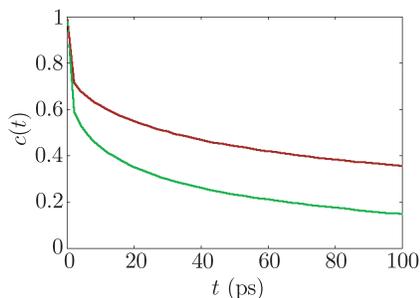


Figure 22. Hydrogen bond correlation function for mutant chain with urea (green) and mutant chain with hydroxyectoine (red) in water at 300 K

Similar simulations with urea and hydroxyectoine were performed for the mutant AFP. Figure 22 shows hydrogen bond lifetime plots for water molecules around the (*ala-ala-ala*) chain in aqueous hydroxyectoine (red) and aqueous urea (green). The hydrogen bond lifetime of water molecules around the mutant chain was greater in the aqueous hydroxyectoine than aqueous urea. However, the antifreeze activity is much smaller compared to the AFP chain (not shown).

This behavior is similar to the behavior of the AFP and the mutant chain in water as discussed previously, as in Figure 13. Also Table 3 depicts the values of rate constant K_f which is a measure of hydrogen bond lifetimes. The rate constant value is very low for the AFP chain in aqueous hydroxyectoine compared to aqueous urea and water. This supports the previous results obtained, predicting an increased antifreeze activity for the AFP chain in aqueous hydroxyectoine. Also literature studies show that the presence of the hydroxyl and carbonyl group enhances the antifreeze activity, showing an increased value of thermal hysteresis [17]. Urea has a hydroxyl group and the bulkier hydroxyectoine possesses both carboxyl and hydroxyl groups, thus, this could be a major contributing factor to an increase in thermal hysteresis.

3.3.1. Influence of urea on protein stabilization

The mechanism of urea interaction with protein has been a major research in the field of computational biochemistry. Despite extensive studies it has been difficult to predict the mechanism of interaction of urea with protein. Insights into the action of urea come largely from experiments that measure transfer free energies of amino acid chains and the peptide backbone. Based on these experiments two different mechanisms have been proposed, an indirect mechanism in which urea is presumed to disrupt the structure of water, thus making hydrophobic groups more readily solvated and a direct mechanism in which urea interacts either directly with the protein backbone, via hydrogen bonds and other electrostatic interactions. Even in the direct mechanism there is a controversy which of the forces are dominant. The direct electrostatic mechanism suggests that urea directly interacts with the protein backbone, via hydrogen bonds and other electrostatic interactions with the polar and charged side chains predominantly [25].

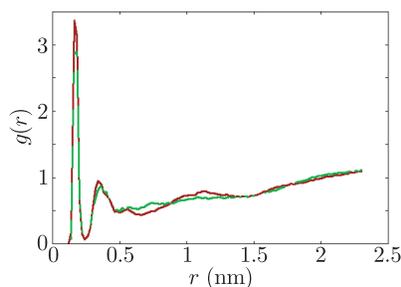


Figure 23. Probability density function between backbone amide hydrogen and oxygen atom of water for AFP in the presence of urea for 10–15 ns (green) and 15–20 ns (red)

Figure 23 shows the pair correlation function of the hydrogen atom of the amide backbone and the oxygen atom of the water molecule for AFP in aqueous urea in different time intervals. The result shows that there is no large deviation of water molecules around the AFP chain, thus, there is no increase in the hydration shell. Simultaneously, it is seen from Figure 24 which depicts the pair correlation

function of amide hydrogen of protein chain and carbonyl oxygen of urea molecule at different time intervals that there is a considerable increase in the presence of urea near the protein chain (11–19). The magnitude of this peak of $g(r)$ at different time intervals as shown in the figure fluctuates around 11–19 during the last 10 ns. This suggests that there are conformational changes introduced in the protein chain due to urea [23].

There are many ordered urea oxygen atoms near the protein amide hydrogen atom in the last 10 ns, but almost no change is found for water oxygen atoms around the amide hydrogen atom of protein.

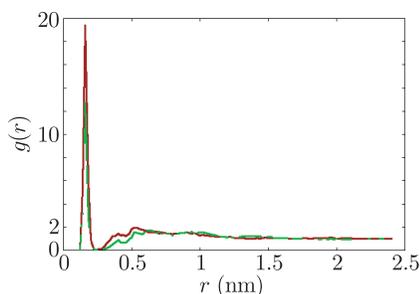


Figure 24. Probability density function between backbone amide hydrogen and carbonyl oxygen atom of urea for AFP in 10–15 ns (green) and last 15–20 ns (red)

The preferential binding parameter is calculated using Kirkwood-Buff integrals as mentioned previously. The probability density function for the protein chain with urea and water was calculated to obtain the preferential binding parameter Γ . Figure 25 clearly shows that urea preferentially binds (positive value of PB coefficient) to the hydrophobic alanine base, the preferential binding parameter is positive as shown in the figure. Figure 26 shows that urea preferentially excludes with the hydrophilic, polar threonine base. However, for $r = 0.5$ nm, the preferential binding parameter value is zero, which frames the first hydration protein shell. This result shows that urea tends to get attracted toward hydrophobic nonpolar alanine residue and tends to bind to it preferentially compared to the threonine residue. Thus, compiling these results we predict that urea forms hydrogen bonds more tightly with protein than water molecules.

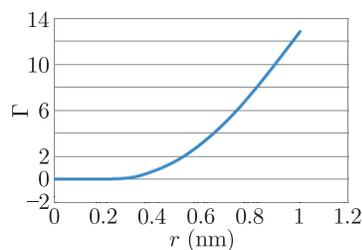


Figure 25. Preferential binding plot of urea with alanine

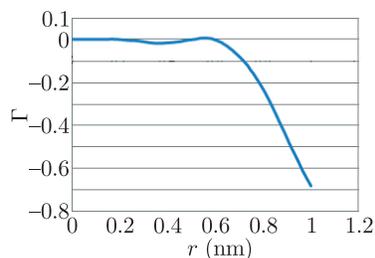


Figure 26. Preferential binding plot of urea with threonine

The preferential binding of urea towards the AFP chain is clearly depicted in Figure 27. This preferential binding of urea to the protein is the primary mechanism by which urea disrupts the native intra-protein hydrogen bonds (more precisely alanine residue) and replaces the water molecules surrounding the protein chain. It is known that the solubility of isolated hydrophobic residues is increased in aqueous urea solution [29], hence, the presence of this nonpolar residues may account for stabilization or destabilization of protein. However, the discussion regarding destabilization of protein in aqueous urea is still unclear. This is studied by calculating the solvent accessible surface area and hydrogen bond density and radius of gyration.

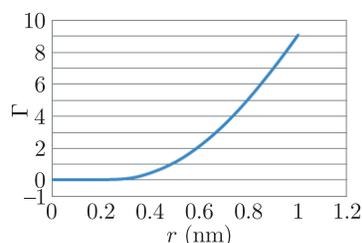


Figure 27. Preferential binding plot of urea with AFP chain

Now, we have shown that urea preferentially binds to the peptide chain, we have also predicted that some of the urea molecules tend to be in close proximity with hydrophobic residues. Figure 28 depicts the number of hydrogen bonds as a function of time for the backbone-backbone (red) bond, backbone-water (green), and backbone-urea (blue) bond. The intra-peptide bond remains constant throughout the course of simulation. The hydrogen bond for urea and protein chain increases from 25–55 in the initial time scale where the number of hydrogen bonds between water and protein chain depreciates significantly (the results are analyzed for the last 15 ns of the simulation run). During the increase in the hydrogen bonds (40–60) between urea and protein, there is a corresponding hike in the radius of the gyration value as shown in Figure 29. The radius of gyration reaches a value of 1.04 nm when the number of hydrogen bonds between urea and protein reaches its peak. This is also supported by a corresponding decrease in the number of hydrogen bonds between water and protein. When

there is a large increase in hydrogen bonds between urea and protein the number of hydrogen bonds between water and protein remains almost constant (Figure 28). Also in Table 4 the hydrogen bond density of water molecules around protein in aqueous urea (6M) shows a lower value compared to aqueous hydroxyectoine, though the total solvent accessible surface area is large compared to aqueous hydroxyectoine. Hence, from these parameters we propose that some of the water molecules near the protein surface should be considerably replaced by urea molecules, thus showing that urea penetrates the protein first, and disrupting the structure sufficiently, creates water frames around proteins.

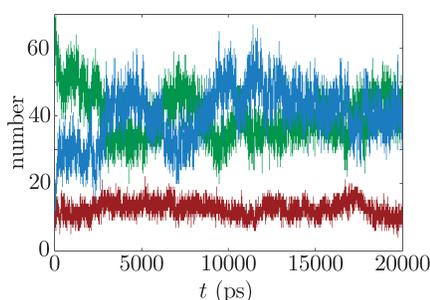


Figure 28. Number of hydrogen bonds as a function of time for AFP in presence of urea water solution, protein-protein bond is shown in red, protein-water in green, protein-urea in blue

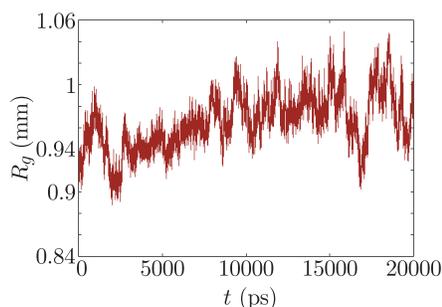


Figure 29. Radius of gyration of AFP chain in aqueous urea

These results suggest a two stage kinetic mechanism for the action of urea. In the first stage, an increase in the urea-backbone interaction results in the swelling of protein, thus, explaining the destabilization of protein chain in the presence of aqueous urea (also supported by a corresponding increase in the solvent accessible surface area, discussed later). Subsequently, water also invades the protein interior although its density around the protein is diminished relative to the bulk value (Figure 28). Hence, urea makes favorable contacts with the hydrophobic residues and seems to enter the protein before water on the exposure of hydrophobic residues. Urea molecules can act as a surfactant to solubilize the protein side chains in water [23].

3.3.2. Influence of hydroxyectoine on protein stabilization

Protecting osmolytes or osmoprotectants are organic molecules that stabilize the folded state of the protein. Here, we study the effect of hydroxyectoine (6M) on protein stabilization. Hydroxyectoine (zwitterionic) is bulkier in comparison with urea, it contains both carboxyl and hydroxyl side chains in comparison with urea. Figure 30 shows the pair correlation plot between backbone amide hydrogen and the oxygen atom of water for AFP in the presence of hydroxyectoine for 10–15 ns (green) and last 15–20 ns (red). The figure depicts a large variation in the number of water molecules surrounding the protein surface in the last 10 ns. The pair correlation value fluctuates from (0.4–1.2). This is in complete contrast with protein in the presence of aqueous urea where there is no profound change in water surrounding the protein surface.

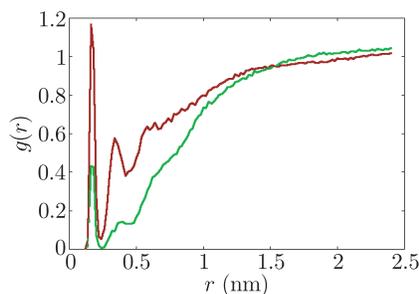


Figure 30. Probability density function between backbone amide hydrogen and oxygen atom of water for AFP in presence of hydroxyectoine for 10–15 ns (green) and last 15–20 ns (red)

Figure 31 shows that hydroxyectoine molecule preferentially excludes (negative value of PB coefficient) the hydrophilic (threonine) residues in the region of 0.4–0.6 nm. Similar effects were found for alanine in aqueous hydroxyectoine. Hydroxyectoine molecules tend to preferentially exclude from the apolar alanine chain as shown in Figure 32.

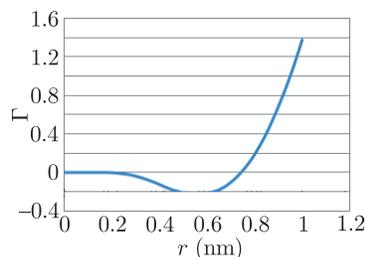


Figure 31. Preferential binding of hydroxyectoine with threonine

This shows a clear preferential exclusion mechanism to the protein backbone. Previous experimental and theoretical studies show that hydroxyectoine molecules tend to bind more to water molecules in comparison with urea [24],

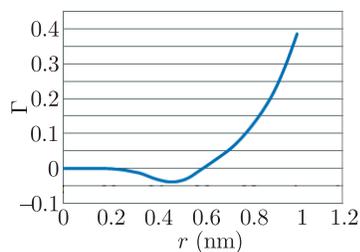


Figure 32. Preferential binding of hydroxyectoine with alanine

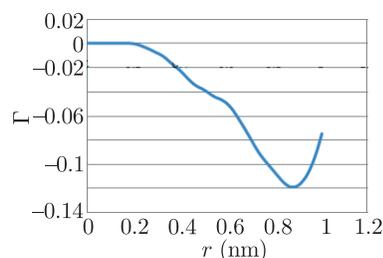


Figure 33. Preferential binding plot of hydroxyectoine with AFP chain

hence they interact strongly with water molecules, enhancing the water structure.

Figure 33 shows a clear trend of preferential exclusion of hydroxyectoine from the AFP chain. Figure 34 depicts the number of hydrogen bonds as a function of time for the backbone-backbone (red) bond, backbone-water (green), and backbone-hydroxyectoine (blue) bond. Even in this the intrabackbone hydrogen bond remains constant throughout the simulation run. The number of hydrogen bonds formed between water and protein remains dominant around 50 compared hydroxyectoines which is around 30. The hydrogen bond density values for water molecules bonded to protein is represented in Table 4, which shows large hydrogen bond densities in comparison to protein in aqueous urea. Also the total solvent accessible surface area is low in comparison to urea.

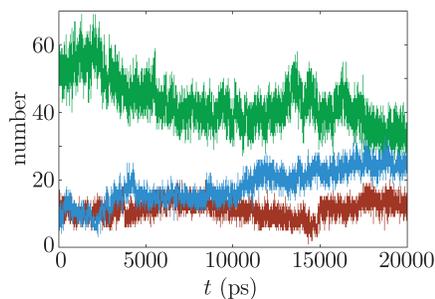


Figure 34. Number of hydrogen bonds as function of time for AFP in presence of HE water solution, protein-protein bond is shown in red, protein-water in green, protein-hydroxyectoine in blue

Table 4. Hydrogen bonding parameters for water molecules around protein chain in presence of cosolutes

Peptide chain	n_{HB}	σ_t (nm ²)	ρ_{HB} Hydrogen bond density (nm ⁻²)
AFP in aqueous hydroxyectoine	42.79	26.360	1.6232
AFP in aqueous urea	40.00	27.625	1.4479

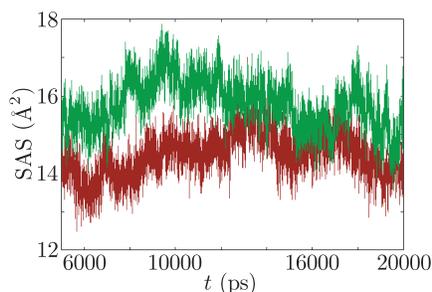


Figure 35. Solvent accessible surface area AFP in presence of urea (green) and hydroxyectoine (red)

Hence, from the above results we propose that the hydroxyectoine molecule tends to enhance the presence of water near the protein surface. The stabilization of protein is predicted by plotting SAS and the radius of gyration plots.

Figure 35 shows the solvent accessible surface area of protein in aqueous urea and aqueous hydroxyectoine which clearly shows that protein solvates more in aqueous urea than in aqueous hydroxyectoine. Also Figure 36 shows that protein tends to remain compact throughout the course of simulation. The value of radius of gyration tends to fluctuate around 0.88–0.91 nm.

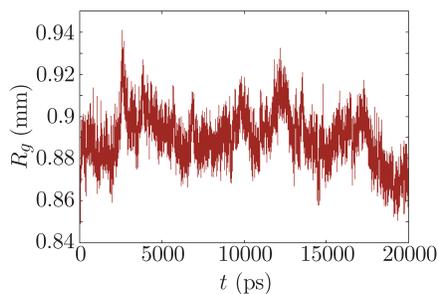


Figure 36. Radius of gyration of AFP chain in aqueous hydroxyectoine

Hence, we conclude from the above results that hydroxyectoine tends to preferentially exclude from the protein chain and enhance the water structure near the protein surface. Thus, an indirect mechanism is proposed. Also the preferential exclusion of hydroxyectoine from the apolar side chain can contribute significantly to an increase in the protein stability [47].

Figure 37 represents the dipole autocorrelation function of water molecules around protein in the presence of aqueous urea and aqueous hydroxyectoine. The results obtained were fitted using an exponential decay curve $C(t) = A \propto \exp(-t/\tau)$, where A represents the percentage of water dipoles reorienting with an average time constant τ .

The result shows a value of $A = 0.5443$ for protein in the presence of hydroxyectoine and $A = 0.9346$ in the presence of urea. This low value of A (percentage of free orienting water dipoles) for hydroxyectoine can be correlated with a decrease in the number of free orienting dipoles of water in hydroxyectoine than in urea. This suggests that more water molecules are bound to the protein in the presence of hydroxyectoine than in urea [48]. This result is shown in Figure 36.

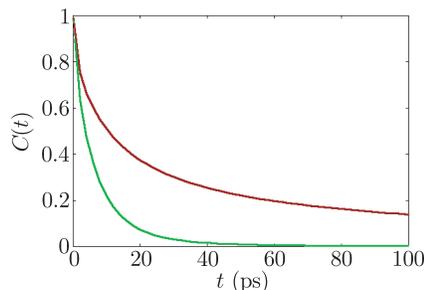


Figure 37. Dipole autocorrelation function for the protein chain in the presence of hydroxyectoine (red) and urea (green)

Table 5 shows the values of time constant τ and percentage free orienting dipoles A . The values of time constant for water molecules around protein in aqueous urea and water are almost similar, but there is a huge increase in the value of τ for water molecules around the protein chain in aqueous hydroxyectoine, showing that water molecules tend to bind more to proteins in the presence of aqueous hydroxyectoine than urea [42]. This result further supports the indirect mechanism of protein stability in aqueous hydroxyectoine.

Table 5. Dipole autocorrelation function for water molecules

Peptide chain	A	τ
AFP in water	0.9491	6.6417
AFP in aqueous urea	0.9346	6.6925
AFP in aqueous hydroxyectoine	0.544	79.2518

3.4. Thermodynamic study of AFP chain in water, aqueous urea and aqueous hydroxyectoine

In this section we discuss the solvation dynamics of apolar alanine and polar threonine residues. The solvation free energy of a solute can be expressed as the

difference between free energies of the solute in solution and as an ideal gas, that is:

$$\Delta A_{sol} = A_{sol} - A_{gas} \tag{35}$$

Here, the Helmholtz free energy appears since the system was simulated under constant volume conditions. To obtain a reversible path between the gas phase molecule to that in solution, a hybrid potential is framed:

$$U(\lambda) = U_{11} + U_{22} + (1 - \lambda)U_{12} \tag{36}$$

where U_{11} and U_{12} represent the potential functions of the solvent molecules ($i = 1$) (including their interaction with each other) and the solute ($i = 2$). The term U_{12} corresponds to the nonbonded interactions between the solute and the solvent and λ is a coupling parameter, which varies from that describing the solute in solution ($\lambda = 0$) to that in the absence of solvent ($\lambda = 1$).

Using the thermodynamic integration method, the solvation free energy is calculated as [49]:

$$\Delta A = \int_0^1 \left\langle \frac{\partial U}{\partial \lambda} \right\rangle d\lambda \tag{37}$$

Proteins are seldom solvated by pure water. Other solvent components, such as buffer salts and stabilizers, are ubiquitous in the laboratory and in formulations of therapeutic proteins. Generally water is a poor solvent for proteins. The presence of these other components, “cosolvents”, generally alters the protein equilibria and the reaction kinetics by perturbing the chemical potential of the protein system. Cosolvents perturb the chemical potential of the protein system by associating either more strongly or more weakly with the protein than water [32].

From previous experimental and theoretical studies it is noted that the apolar alanine residue has a solvation free energy of 2.27 kcal/mol, whereas polar threonine has a solvation free energy of -4.37 kcal/mol in the SPC/E water model [50]. Thus, it can be noted that the solubility of a threonine residue in the water model is more favorable than alanine. This may be due to the presence the of hydrophobic methyl side chain in the alanine residue.

Table 6. Transfer free energy calculations

Peptide chain	Γ (PB parameter)	μ_{tr} , transfer free energy (kJ/mol) (for $r < 0.7$ nm)
AFP in aqueous hydroxyectoine	-0.04736	0.118
AFP in aqueous urea	2.360	-5.89

Transfer models are used as discussed in Section 2 to identify the stability and influence of cosolvents on the stability of proteins. $\Delta\mu_{tr}$ depicts the transfer free energy of protein from pure water into a mixed solvent system. It follows from the transfer model analysis that the transfer of an alanine and threonine residue from pure water to aqueous urea is favorable in the preferential binding regime (< 1) nm. Whereas the transfer of free energy is unfavorable in the case of

aqueous hydroxyectoine. Thus, we can predict that the AFP chain is more stable in aqueous hydroxyectoine than in pure water, while the protein tends to lose its stability in the presence of aqueous urea enhancing the solubility of amino acid chains. This discussion supports the previous results for the direct mechanism of destabilization for the protein chain in urea and an indirect mechanism of stabilization for the protein chain in hydroxyectoine as cosolvents [35].

4. Conclusions

The study of antifreeze protein and its dynamics in water has been an interesting field of research for the past four decades. Understanding the mechanism of freeze resistance of an AFP has attracted much interest in the literature. Previous studies show the adsorption inhibition mechanism as the dominant mechanism, however, mutagenesis experiments show that it is also the presence of hydrophobic groups accounts for the binding [2]. In this work a theoretical study of antifreeze proteins in the presence of the SPC/E water model was studied. The hydration dynamics results revealed that the retarded water dynamics in the binding plane of proteins and heterogeneous distribution of lifetimes for different planes in protein was responsible for the antifreeze activity [7, 18]. The influence of osmolytes on enhancement of antifreeze activity was demonstrated, as well. Hence, we obtain complete validation of the experimental results studied by Havenith and Leitner. Thus, we propose that the retarded water dynamics of water molecules around protein is responsible for the antifreeze activity [7]. The relation between retarded water dynamics and thermal hysteresis is not well established with the obtained results, whereas it follows from the experimental background of the Havenith study that long range retarded dynamics is responsible for the thermal hysteresis activity.

The control of the protein stability by addition of cosolvents is a biophysical problem of great fundamental interest, with many application in biotechnology. The mechanism by which urea destabilizes protein has been a field of research for a long time. Here, we used a simulation technique to study the effects of the AFP chain in aqueous urea (6M). The swelling of protein was absorbed in the presence of aqueous urea. Whereas simulations were performed for protein chain in aqueous hydroxyectoine (6M) to study the effect of protein stabilization. An indirect mechanism by which hydroxyectoines enhance water structuring around the protein chain was proposed. The thermodynamic results were quoted to support the discussion provided in the results. A transfer model studies were performed in order to support the mechanism proposed for proteins in aqueous urea and hydroxyectoine.

Acknowledgements

I would like to acknowledge all the members in Professor Christian Holm's group. In particular, I greatly appreciate the help and kindness of my advisor, Jens Smiatek, PhD. He has given me much invaluable inspiration for my thesis. I would also like to thank the members of Institute of Computational Physics for

their support during my stay in this institute. I would also like to thank Professor J. Rybicki for his guidance with my thesis. I am also thankful to all my friends and colleagues who supported me in the course of writing my thesis.

References

- [1] www.rcsb.org/pdb/101/motm.do?momID=120
- [2] De Vries A L and Wohlschlag D E 1969 *Science* **163** 1073
- [3] Jorov A, Zhorov B S and Yang D S 2004 *Protein Sci.* **13** (6) 1524
- [4] Smiatek J, Harishchandra R K, Rubner O, Galla H-J and Heuer A 2012 *Biophys. Chem.* **160** 62
- [5] Collins K D, Neilson G W and Enderby J E 2007 *Biophys. Chem.* **128** 95
- [6] Marshall C B, Daley M E, Graham L A, Skyes B D and Davies P L 2002 *FEBS Lett.* **529** (2-3) 261
- [7] Ebbinghaus S, Meister K, Brn B, De Vries A L, Gruebele M and Havenith M 2010 *JACS Chem.* **132** 12210
- [8] Fletcher G L, Hew C L and Davies P L 2001 *Annu. Rev. Physiol.* **63** 359
- [9] Capicciotti C J, Doshi M and Ben R N 2013 *Ice Recrystallization Inhibitors: From Biological Antifreezes to Small Molecules*, Department of Chemistry, D'Iorio Hall, University of Ottawa, Ottawa, ON, Canada
- [10] Raymond J A and De Vries A L 1977 *Proc. Natl. Acad. Sci. USA* **74** (6) 2589
- [11] Schüttelkopf A W and van Aalten D M F 2004 *Acta Crystal.* **D60** 1355
- [12] Greiner W, Neise L and Stöcker H 2001 *Thermodynamics and Statistical Mechanics*, Springer
- [13] Hunenberger P H 2005 *Adv. Polymer. Sci.* **173** 105
- [14] Berendsen H J C, Postma J P M, van Gunsteren W F, DiNola A and Haak J R 1984 *J. Chem. Phys.* **81** (8) 3684
- [15] Karatsas I and Shreve S 1997 *Brownian Motion and Stochastic Calculus*, 2nd Ed., New York, Springer-Verlag
- [16] Hess B, Kutzner C, van der Spoel D and Lindahl E 2008 *J. Chem. Theory Comput.* **4** (3) 435
- [17] Amornwittawat N, Wang S and Banatiao J 2009 *Biochim. Biophys. Acta* **1794** 341
- [18] Meister K, Ebbinghaus S, Xu Y, Duman J G, De Vries A L, Gruebele M, Leitner D M and Havenith M 2012 *Proc. Natl. Acad. Sci. USA*, doi:10.1073/pnas.1214911110
- [19] Makhatadze G I and Privalov P L 1995 *Adv. Protein Chem.* **47** 307
- [20] Marcus Y 2009 *Chem. Rev.* **109** 1346
- [21] Wiggins P M 2001 *Cell. Mol. Biol.* **47** 735
- [22] De Xammar Oro J R 2001 *J. Biol. Phys.* **27** 73
- [23] Hua L, Zhou R, Thirumalai D and Berne B J 2008 *Proc. Natl. Acad. Sci. USA* **105** (44) 16928
- [24] Smiatek J, Harishchandra R K, Rubner O, Galla H-J and Heuer A 2012 *Biophys. Chem.* **160** 62
- [25] Robinson D and Jencks W 1965 *J. Am. Chem. Soc.* **87** 2462
- [26] Arakawa T and Timasheff S N 1985 *Biophys. J.* **47** 411
- [27] Timasheff S N 2002 *Biochemistry* **41** 13473
- [28] Lemkul J A, Allen W J and Bevan D R 2010 *J. Chem. Inf. Model.* **50** (12) 2221
- [29] Luzar A and Chandler D 1996 *Hydrogen Bond Kinetics in Liquid Water*, Department of Chemistry, University of California, USA
- [30] Luzar A 2000 *J. Chem. Phys.* **113** 10663
- [31] Yu I, Jindo Y and Nagaoka M 2007 *J. Phys. Chem.* **B111** 10231
- [32] Baynes B M, Wang D I C and Trout B L 2004 *Proteins in Mixed Solvents: A Molecular-Level Perspective*, Department of Chemical Engineering, M.I.T., Cambridge, USA

- [33] Kirkwood J G and Buff F P 1951 *J. Chem. Phys.* **19** 774
- [34] Ben-Naim A 1992 *Statistical Thermodynamics for Chemists and Biochemists*, Plenum Press, New York
- [35] Canchi D R and Garcia A E 2013 *Annu. Rev. Phys. Chem.* **64** 273
- [36] Bencivenga F, Cimatoribus A, Gessini A, Izzo M G and Masciovecchio C 2009 *J. Chem. Phys.* **131** 1445021
- [37] Richmond T J 1984 *J. Mol. Biol.* **178** 63
- [38] Eisenhaber F, Lijnzaad P, Argos P, Sander C and Scharf M 1995 *J. Comp. Chem.* **16** 273
- [39] Ebbinghaus S, Kim S J, Heyden M, Yu X, Heugen U, Gruebele M, Leitner D M and Havenith M 2007 *Proc. Natl. Acad. Sci. USA* **104** (52) 20749
- [40] Nicodemus-Johnson J, Silic S, Ghigliotti L, Pisano E and Cheng C-H C 2011 *Genomics* **98** 194
- [41] Zangi R, Zhou R and Berne B J 2009 *J. Am. Chem. Soc.* **131** (4) 1535
- [42] Luzar A and Chandler D 1996 *Nature* **379** 55
- [43] Li N, Andorfer C and Duman J 1998 *J. Exp. Biol.* **201** 2243
- [44] Amornwittawat N, Wang S, Duman J G and Wen X 2008 *Biochim. Biophys. Acta* **1784** 1942
- [45] Jorov A, Zhorov B S and Yang D S C 2004 *Protein Sci.* **13** 1524
- [46] Raymond J A and De Vries A L 1977 *Proc. Natl. Acad. Sci. USA* **74** 2589
- [47] Stanley C and Rau D C 2008 *Biochemistry* **47** 6711
- [48] Balasubramaniam S, Bandyopadhyay S, Pal S and Bagchi B 2003 *Current Science* **85** 1571
- [49] Wan S, Stote R H and Karplus M 2004 *J. Chem. Phys.* **121** (19) 9539
- [50] Shirts M R, Pitner J W, Swope W C and Pande V S 2003 *J. Chem. Phys.* **119** 5740