Hydration layer dynamics and association mechanisms of food and antifreeze proteins

A Molecular Dynamics and Transition Path Sampling study

Brotzakis, Z.F.

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
Other

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Chapter 1

Introduction

1.1 Proteins in nature and technology

1.1.1 Protein function

Proteins are the workhorses of the living cells. They consist of linear polypeptide chains made out of amino acids linked together through covalent peptide bonds. In the cell the processes of transcription and translation convert DNA to mRNA, which subsequently serves as a template for the production of amino acid chains in the ribosomes [1]. Interestingly, there are 20,000-23,000 human genes that encode proteins, which with the above procedure, may produce hundreds of thousands distinct proteins of different functionality when expressed.

Proteins differ in amino acid sequence, structure, shape, and size, thus giving rise to different biological functions, including enzyme catalysis (enzymes), transport of ions through membranes (membrane transport proteins), regulation of cell activity by e.g. transmitting extracellular signals to the cell’s interior (signaling proteins), and translation of other proteins or organelles (molecular motors). An important point is that proteins are functional under physiological conditions at which they fold into a minimum free energy 3D configuration, their native state. Proteins with common sequence are likely to belong to a common ancestor (homologs), leading to similarities in structure, e.g. the similar X-ray crystallographic structure and similar sequence (43%) of myoglobin and hemoglobin. Notwithstanding this connection between sequence structure and function, there is still not a clear relation between the sequence, structure and functionality. Moreover, the mechanism of protein folding has not been completely understood, especially for larger proteins. Answering these questions would help de-
signing novel proteins of tailored stability and dynamics in applications such as vaccines, drug delivery vesicles, new enzymes etc. The folding process is not 100 percent accurate, and misfolding and/or aggregation of proteins can occur. This can lead to neurodegenerative diseases like Alzheimer’s, Parkinson’s, Huntington, type 2 diabetes, Bovine Spongiform Encephalopathy (BSE) [2–4]

1.1.2 Food Proteins

In this thesis we are mostly interested in proteins encountered in food applications. Examples of protein rich food systems are soy, pea, gelatin, but also milk and dairy products. In particular, we will mostly look at whey proteins, which are a substantial component of milk. While clearly originating from biological resources (plants, fungi, algae and animals), food proteins are exposed to a much higher range of temperature (-20 to 100 °C), pH (1-10) and salinity conditions, related to food preparation and application. High temperature or addition of co-solutes can cause protein denaturation, which usually leads to aggregation. Food proteins, depending on the application, encounter many different types of aggregates, such as 3D networks of aggregates spanning discrete space regions or even the entire space in the case of percolation of gels. Food proteins can be used as surfactants in order to stabilize the interfaces of heterogeneous food materials, e.g. foams and oil-water interfaces. Notably, protein-rich food systems hardly ever consist of proteins alone, but often contain other hydrocolloids or surface active molecules, controlling their aggregation [5].

1.1.3 Ice binding and anti-freeze proteins

Ice Binding Proteins (IBPs) occur in a variety of organisms and have a variety of functions, such as blocking the growth of ice crystals, structuring ice [6, 7], inhibiting ice recrystallization [8], and promoting adhesion [9]. Their unique capability to regulate the ice crystal growth makes them great candidates for use in many applications in food technology, medicine and material science.

For many freeze-avoiding species such as fish, insects etc, IBPs prevent the growth of ice in their blood stream [10]. This subclass of IBPs is called anti-freeze proteins (AFPs) as they predominantly lower the freezing point of ice below its melting point in a non-colligative manner. For instance, fish AFPs show a depression of the freezing point by ≈ 2°C at serum levels of 30-40 mg ml⁻¹. Anti-freeze proteins have various shapes and sizes and are grouped in fish and non fish AFP categories (see Fig. 1.1.1). Fish AFPs are divided into type I,II,III and IV, with type I having a repetitive aminoacid sequence and type II,III having a non-repetitive sequence and globular structure. In this thesis we will investigate solvation properties of anti-freeze
protein of type $III$, as well as a new synthetic cyclic peptide that self-assembles into an anti-freeze nanotube structure.

1.2 Protein structure and interactions

1.2.1 Structure of proteins

Apart from the primary structure-sequence of amino acids mentioned above, proteins exhibit secondary, tertiary and quaternary structure. Secondary structures are spatial arrangements of protein segments such as $\alpha$-helices, $\beta$-sheets and $\beta$-turns, which involve mainly backbone hydrogen bonds. The tertiary structure relates to the total 3D conformation of a folded polypeptide chain. Contrary to the secondary structure, the tertiary structure is stabilized by hydrophobic interactions, salt bridges, side chain hydrogen bonds and disulfide bridges. These interactions further stabilize the secondary structure. Finally, the last level of structural categorization is the quaternary structure. As proteins can be multimeric, the relative positions of the different domains form the quaternary structure (e.g. Hemoglobin).
1.2.2 Protein-protein interactions

**Electrostatic repulsion and Van der Waals attraction: DLVO theory**

Native globular proteins in solution can be viewed as colloidal particles suspended in an electrolyte solution. While in principle the underlying atomistic interactions govern the behaviour of such suspensions, it is possible to describe the interaction between such particles in terms of effective potentials between the centers of mass, by integrating out the microscopic degrees of freedom. While most of the interactions are of electrostatic nature, it is convenient to classify the effective interactions in different categories: screened electrostatic interactions, hydrophobic interactions, Van der Waals interaction, hydration interactions, and remaining molecular interactions such as salt bridges and hydrogen bonds.

Electrostatic and Van der Waals interactions are often treated together in the so called DLVO theory. Here the particles interact via a sum of two body interactions: of the repulsive screened electrostatic and the Van der Waals attraction. For spherical particles (e.g. globular proteins) the interaction energy $W$ becomes [12]:

$$W = \frac{1}{2}ZRe^{-\kappa D} - \frac{AR}{12D}, \quad (1.1)$$
where \( Z \) is the interaction constant (see eq. 1.2), \( R \) is the sphere radius, \( D \) is the distance between the two spheres, \( A \) is the Hamaker constant and \( \kappa^{-1} \) is the Debye length. The interaction constant \( Z \) is given by eq 1.2,

\[
Z = 64\pi\varepsilon_0\varepsilon \left( \frac{k_BT}{e} \right)^2 \tanh^2 \left( \frac{ze\psi_0}{4k_BT} \right),
\]

(1.2)

where \( \varepsilon_0 \) is the dielectric permittivity in vacuum, \( \varepsilon \) is the dielectric constant of water, \( e \) is the elementary charge, \( T \) is the temperature and \( k_B \) the Boltzmann constant, \( z \) is the valence of an ion in the electrolyte solution and \( \psi_0 \) is the surface potential. The charges on the sphere attract counter ions from the solution, creating a double layer, which screens the electrostatic interaction. Note that eqs. 1.2, and the first term of 1.1 hold for symmetric electrolyte solutions (e.g. 1:1, 2:2). The Debye length is related to the electrostatic screening and depends on the ionic concentration as follows (eq 1.3),

\[
\kappa^{-1} = \sqrt{\frac{\varepsilon_0\varepsilon k_BT}{2N_Ae^2I}},
\]

(1.3)

where \( I \) stands for the ionic strength, and \( N_A \) is Avogadro’s number. DLVO interactions scale with the radius of the interacting bodies. In Fig. 1.2.2 it is assumed that
particles have radius $R=100$ nm. Globular proteins of moderate size do not exceed a radius of a few nanometres, therefore the conclusions mentioned below are most evident in much larger size proteins or colloids. For low ionic strength (dilute electrolyte solutions) the Debye length increases, giving rise to an effective electrostatic repulsion, which forms a large energy barrier between 1 and 5 nm (see Fig. 1.2.2). Upon increasing the ionic concentration, the Debye length is decreased, therefore screening the electrostatic repulsion and reducing the barrier, while forming a secondary minimum. At very high salt concentration the effective screening is so large that the only force into play is the Van der Waals attraction. Increasing the salt concentration is therefore likely to destabilize the protein solution. Also changing the surface charge, e.g. by changing the pH, can induce aggregation. Through a change in protonation state of side chains the protein can reach its isoelectric point, where the electrostatic repulsion becomes minimal.

**Hydrophobicity**

Hydrophobic interaction is very common in biology and drives many processes such as folding and hydrophobic association. Water in the hydrogen bond network around hydrophobes, although more coordinated, is entropically hindered to restructure the network of hydrogen bonds [13]. However, for bigger hydrophobic objects the formation of a complete hydrogen bond network is geometrically impossible making hydrophobic hydration energetically and entropically unfavourable [14]. For spheres at a critical distance, this effect becomes prohibiting, and water at the interface evaporates. This gives rise to a dewetting transition [15, 16] and therefore to an effective attraction between the hydrophobic spheres.

**Other surface interactions**

Besides Van der Waals, screened electrostatic interactions, and hydrophobic attractions, there exist other non-DLVO forces that are present on protein association. The most obvious one is the direct hydrogen bonding between residues of different proteins. Further, salt bridges provide strong electrostatic interactions between proteins, thus providing a highly stabilizing role. Also, hydrophilic protein surfaces bind water hydration layers. When two proteins are approaching each other, the hydration layers of the two feel each other and cause repulsive hydration forces due to the change in the structuring of each individual layer. These hydrophilic hydration repulsion forces are oscillatory [12, 17, 18].

Other surface interactions that are crucial for protein association include water bridging interactions, where a water makes a double hydrogen bond with the protein. A
thermodynamic model by Ben-Naim predicts the importance of bridging waters in reducing the free energy of binding and driving the assembly towards tighter binding [19].

1.3 The role of solvent (water) and co-solutes (e.g. salts) on protein stability

The solvent mediates protein interactions as well as being an agent on its own, driving protein motion (hydration forces, H-bond network, and the hydrophobic/hydrophilic effect). Below, we briefly introduce the importance of water, its relevance as a solvent around proteins, and its relation to salts. Ions are very important in biology. Under physiological salt conditions (∼100 mM), ions participate in the osmotic stress regulation and other main processes [1]. In food related applications ions are used to alter the stability of colloids or proteins towards preferred phases (e.g by reducing protein-protein repulsion). Water by itself exhibits unique properties many of which can be attributed to the unique tetrahedral H-bond structure that is prevalent both in liquid and crystalline water. The water H-bond structure and dynamics is altered by the presence of a surface. In this thesis we investigate how the H-bond structure and dynamics is influenced by the presence of a protein surface.

Importance of water around proteins (hydration layers)

Water around proteins plays a significant role in many processes, from protein folding to enzyme function. Hydrophobic interactions are known to drive proteins to fold in order to increase the translational and orientational entropy of water [20, 21]. Moreover, water mediated hydrogen bond interactions assist in the stabilization of the secondary and tertiary structure [1]. Interestingly, proteins slow down the hydration water translation and orientation by a factor ∼ 3-5. Finally, the slow dynamics of water is important for enzymes such as metalloprotease MT1-MMP. Experiments and simulations showed that for this enzyme, the formation of the Michaelis complex, is driven by a slowdown of water reorientation towards the active site [22].

Hofmeister series

Addition of salt can stabilize or destabilize the protein solution. In the 19th century, the Czech protein scientist Hofmeister, investigated the influence of ions in the precipitation of yolk egg protein and other small colloids. The Hofmeister series (see eq. 1.4)
is an ordering of ions according to their propensity (measured by concentration) to precipitate proteins (salting out).

\[
CO_{3}^{2-} > SO_{4}^{2-} > S_{2}O_{3}^{2-} > H_{2}PO_{4}^{-} > F^{-} > Cl^{-} > Br^{-1} \approx NO_{3}^{-} > I^{-} > ClO_{4}^{-} > SCN^{-} \quad (1.4)
\]

The origin of the Hofmeister series is not well understood and has been attributed to a number of causes. Jones and Dole [23] have shown that some ions affect water viscosity positively (kosmotropes) and others negatively (chaotropes). In other words, ions are structuring or disrupting the H-bond network of electrolyte solutions.

On the other hand, there has been an increasing number of studies suggesting the ion-protein interactions causing the precipitation/unfolding propensity, however without yet having reached a consensus on the mechanism of the series [24]. However, there is evidence that Hofmeister ions with a higher propensity have a greater affinity for amide groups and are excluded from hydrophobic groups [24].

**Water reorientation dynamics: Jump models**

Water H-bond structure and dynamics around proteins plays an important role in biology, such as in protein folding, ion channels and enzyme reactions. While in the Debye model [25] water motion was believed to be a barrierless random angular diffusion, more recent models indicate that water reorientation and motion is dictated by large-amplitude and sudden angular jumps. Here, we first briefly present some experimental techniques measuring water reorientation dynamics followed by a brief introduction to some of the recent models treating water reorientation and the hydrogen bond exchange mechanism.

IR pump-probe spectroscopy [26–28] has femtosecond time resolution and therefore can capture the very fast water reorientation. However, it suffers from the fact that it cannot address time delays larger than 10 ps. On the other hand nuclear magnetic resonance spectroscopy (NMR) [29, 30] can measure the spin-lattice relaxation time and give information about the average rotational relaxation but not its short time evolution.

One way to describe water reorientation dynamics is by the time correlation function of the OH bond vector. This function is given by eq. 1.5,

\[
C_{n}(t) = \langle P_{n}[\mathbf{u}(0) \cdot \mathbf{u}(t)] \rangle \quad (1.5)
\]

where \( P_{n} \) stands for the \( n^{th} \) order Legendre polynomial and \( \mathbf{u}(t) \) is the OH vector at time \( t \). By plotting the time correlation function \( C_{2} \) (see Fig. 1.3.1) one can extract the timescale of different mechanisms [31–33]. The fast (< 200 fs) librational motion
involves fast wobbling of the OH vector around the axis of the donor-acceptor H-bond (see Fig. 1.3.1a_i,a_ii). A hydrogen bond jump event (see Fig. 1.3.1a_iii and Fig. 1.3.1b) involves the elongation of the initial H-bond, while a new water oxygen acceptor is approaching. Once the two H-bond accepting oxygens are at the same distance, the OH jumps from the old to the new H-bond acceptor. In an effort to quantify the kinetics of the angular mechanism, Ivanov et. al. [32] developed the jump model. This model assumes constant amplitude jumps between hydrogen bonded water with frequency \( 1/\tau_{jump} \).

However, in the absence of librational and frame tumbling motion (see Fig. 1.3.1a_i and Fig. 1.3.1a_iii respectively) in the jump model, an extended version was developed by Laage and Hynes [33], the extended jump model (EJM). There, the overall water reorientation timescale (\( \tau_{EJM} \)) is dictated by a hydrogen bond switching timescale (\( \tau_{jump} \)) and a frame tumbling timescale (\( \tau_{frame} \)), as shown in eq. 1.6.
When water solvates a solute molecule, its reorientation is retarded. In the EJM, this retardation can be attributed to the transition state excluded volume. The excluded volume in concentrated solutions creates convex environments (see Fig. 1.3.1d). The retardation factor due to excluded volumes ($\rho_v$) is larger than two as opposed to convex surfaces of dilute solutions (see Fig. 1.3.1c). While the presence of solutes can cause stronger water-solute H-bonds, the retardation of water reorientation dynamics can also be attributed to the relative difference in H-bond strengths of the pure bulk compared to the solution ($\rho_{HB}$). All the above effects define the overall retardation of the reorientation, shown in eq. 1.7

$$\tau_{EJM} = \frac{1}{\tau_{\text{jump}}} + \frac{1}{\tau_{\text{frame}}},$$  \hspace{1cm} (1.6)$$

$$\tau_{\text{jump}} = \rho_v \rho_{HB} \tau_{\text{jump}}^{\text{bulk}}.$$ \hspace{1cm} (1.7)

### 1.4 Protein association

Association and self-assembly of bio-molecules involves many controlled processes occurring in biology, such as protein association, protein-ligand binding, signalling
and inter-cellular communication. Protein self-assembly and association is a controlled process not only in biology but in technology as well. Depending on the conditions self-assembly can result in a variety of different aggregates, which self-assemble/aggregate with different mechanisms (see Fig. 1.4.1). For instance, by controlling the conditions (e.g. pH) one can change the mechanism of self-assembly, and therefore grow different aggregates (fibrous or worm-like) having different properties (e.g. texture in food industry). Aggregation and fibrilization can also be an unwanted process, such as in the case of amyloidogenic peptides. In this thesis we will investigate how a new synthetic cyclic peptide self-assembles into a nanotube structure.

On a different note, the dimerization transition of native globular proteins is very important on the biological level as it is important in regulation networks as well as the first step of oligomerization, which later forms worm-like fibrils [36]. In food systems (e.g. milk), dimers can occur in thermodynamic equilibrium with monomers [37]. In the following section we introduce the thermodynamics and dynamics of dimer formation.

**Monomer-dimer equilibrium**

From a thermodynamic perspective, the free energy difference between the dimer state (D) and monomer state (M) is given by eq. 1.8.

\[-RT \ln(K) = \Delta G = \Delta H - T \Delta S\]  \hspace{1cm} (1.8)
Here, $K$ stands for the dimerization equilibrium constant, and $\Delta G$, $\Delta H$ and $\Delta S$ are the Gibbs free energy, enthalpy and entropy of dimer formation, respectively. $T$ is the temperature and $R$ is the gas constant. At pH values higher or lower than the isoelectric point $pI$, the electrostatic repulsion favours the monomer population and mainly contributes positively to the enthalpy $\Delta H$ of formation of dimers. Moreover, the reorientation and translation entropy of the proteins is higher in the monomer state, decreasing the $-T\Delta S$. Negative $\Delta G$, where $\Delta G = G_D - G_M$ corresponds to a driving force towards the dimer (see Fig. 1.4.2). When approaching the $pI$, the effective electrostatic screening repulsion becomes smaller, and at the expense of decreasing the translational and orientational entropy, proteins attract each other and the dimer formation is favoured. As mentioned above the effective attraction is dominated by van der Waals interactions, salt bridges, hydrogen bonds (decreasing $\Delta H$), and for some systems by hydration forces.

Protein association (e.g. dimer association) processes are activated processes involving rare transitions between states that are separated by an free energy barrier. The thermodynamic perspective of dimerization was discussed above. Below we discuss reaction rate theories intending to quantify the kinetics of rare transitions, notably phenomenological rate equations. The Arrhenius equation [38] and the transition state theory (TST).

Using a statistical mechanics perspective, dimerization can be pictured as a transition between stable states or free energy minima N (native dimer) and U (unbound monomers), separated by an activation barrier (see Fig. 1.4.3). The rate of the transition between the two minima is inversely proportional to the exponent of the barrier height [39–41]. Consequently, this means that the dwell time in the stable states is much larger than the transition time, $\tau_{state} \gg \tau_{mol}$, the so-called separation of timescales. For a two state activated dynamical process, one can express the kinetics in terms of phenomenological rate equations shown in eqs. 1.9-1.10. The solution of these equations is given in eq. 1.11 and states that a system which has realized a density fluctuation ($\Delta C_N(0)$) of species N from the equilibrium density $\langle C_N \rangle$ at time zero, returns to its the equilibrium population exponentially according to the reaction time $\tau_{rxn}^{-1} = k_{NU} + k_{UN}$, where $k_{NU}$ and $k_{UN}$ are the forward and backward rate constants respectively.

$$\frac{dC_N}{dt} = -k_{UN}C_N(t) + k_{NU}C_U(t)$$ \hspace{1cm} (1.9)

$$\frac{dC_U}{dt} = k_{UN}C_N(t) - k_{NU}C_U(t)$$ \hspace{1cm} (1.10)

$$\Delta C_N(t) = C_N(t) - \langle C_N \rangle = \Delta C_N(0) \exp \left( \frac{-t}{\tau_{rxn}} \right)$$ \hspace{1cm} (1.11)
Figure 1.4.3: Schematic projection of the dissociation free energy of β-lactoglobulin dimers as a function of the reaction coordinate from the native dimer state $N$ to the unbound state $U$. $TS$ stands for the transition state at the top of the barrier.

The Arrhenius equation 1.12 addresses the kinetic rate constants of chemical reactions,

$$k = Ae^{-\frac{E_a}{k_BT}}$$  \hspace{1cm} (1.12)

with $A$ the pre-exponential factor and $E_a$ the activation energy and $k_B$ the Boltzmann constant. The Arrhenius equation expresses the transition rate constant as a function of the temperature, the frequency of collisions $A$ and the probability $e^{-\frac{E_a}{k_BT}}$ to result in a reaction, given a collision. A more fundamental way to look at a reaction rate constant is the Eyring-TST equation [39]. It states that the rate of transformation from one free energy minima to another, e.g. $U$ (unbound state) to $N$ (native dimer) (see Fig. 1.4.3) reduces proportionally to the exponent of the free energy barrier height [39–41], and is given by equation 1.13,

$$k = \frac{k_B T}{h} \kappa e^{-\frac{\Delta G^\ddagger}{k_BT}} = \frac{k_B T}{h} \kappa e^{-\frac{\Delta H^\ddagger}{k_BT}} e^{\frac{\Delta S^\ddagger}{k_B}}$$  \hspace{1cm} (1.13)

with $h$ being the Planck constant, $\kappa$ the transmission coefficient, $\Delta G^\ddagger$, $\Delta H^\ddagger$ and $\Delta S^\ddagger$ the Gibbs free energy, enthalpy and entropy of activation, respectively. The double dagger symbol denotes the values at the transition state. Note that the transition state theory is approximate, and that more elaborate theories or simulations are required to make progress.
The rate of protein association in liquids is bound from above by the diffusion of proteins in the available volume $V$ before they are close enough to associate. Such reactions, where the slow part of the reaction is bringing together the proteins, are denoted as diffusion controlled reactions. The bimolecular association rate constant $k_A$ for two uniformly interacting spherical particles is given by the Smoluchowski rate $k_A = 4\pi DR/V$, where $D = D_A + D_B$ and $R = R_A + R_B$. $D_A$ and $D_B$ are the diffusion constants of particles A and B and $R_A$ and $R_B$ are the radii of particles A and B respectively.

The kinetics and mechanisms of association, even on the dimer level, is poorly understood and varies with the nature of the proteins. Many proteins associate into homodimers with a (experimentally determined) rate constant $k_{on} \sim 10^6 M^{-1}s^{-1}$, which is three orders of magnitude slower than the diffusion limited Smoluchowski association rate constant predicts. This sounds reasonable because association of proteins into stable dimers requires very specific favorable interactions in the native (bound) dimer. However, an estimate of the association rate purely based on the geometric probability that a protein would randomly collide into the correct dimer state would be six orders of magnitude smaller than the diffusion limited rate [42]. Nonspecific interaction can enhance the rate [42, 43] even in the absence of strong hydrophobic and/or charged patches that could steer the assembly. The presence of strong hydrophobic or electrostatic interactions steering the association would result in a much faster association with a rate in the order of $10^9 M^{-1}s^{-1}$, such as in the barnase-barstar complex [44]. In the work by Ahmad et. al. it was shown that for the barnase-barstar complex, electrostatic steering drives the association and causes a very high association constant [44] $k_A = 10^8 - 10^9 M^{-1}s^{-1}$, close to the diffusion limited association rate[42]. Nevertheless, the driving force and kinetic mechanism for protein association, even in the absence of such steering forces, is not well understood.

1.5 Molecular simulations

The laws of nature are expressed in equations which very scarcely can be solved analytically (without the use of a computer), e.g. Newtonian motion of three interacting bodies. Before computers, in order to predict material properties, scientists resolved to approximate theories (e.g VDW equation, Debye Huckel theory). However these theories are approximate and often are valid in few systems, thus lacking generality. Computer simulations/computer experiments serve as predictions which test/explain/validate the theory or the experiment. Here we make use of molecular dynamics simulations using all atom force fields. The MD simulations are limited in time and length scales. Especially for rare events, specific techniques need to be developed. One such technique is
Transition Path Sampling (TPS), which samples rare trajectories committing to stable states. Both MD and TPS are used in this thesis.

1.6 Aim of the thesis

The aim of this thesis is three-fold. Firstly, to better understand the role of water around single proteins in terms of structure and dynamics, secondly to understand the association and self-assembly mechanism of multiple proteins, and thirdly to understand the role of water in the association mechanism. Finally, as protein association is a rare event, we develop a new TPS shooting move which can efficiently sample rare transitions between states, separated by asymmetric barriers.

Understanding water structure and dynamics around proteins

Protein hydration water plays a significant role in determining the dynamics and structure of proteins, in facilitating protein folding, in recognizing ice crystal planes (anti-freeze proteins) and in mediating protein association. In this thesis we address the following research questions:

- What is the H-bond structure and dynamics of water around hydrophilic and hydrophobic groups. How does unfolding affect the overall dynamics of water?

- How does water H-bond structure correlate to water reorientation dynamics at the Ice Binding Site of an anti-freeze protein?

Understanding of dynamical mechanisms of association and self-assembly of multiple proteins, and the role of water in these processes

Protein association, especially in dilute conditions is by itself interesting as its mechanism is yet not clear. Moreover, protein association is pivotal in the first steps of self-assembly. Atomistic insight into the association mechanism and initial steps of self-assembly helps understanding and experimentally control these processes. As these processes are rare events, developing computational methods which can better sample these transitions is very important. In order to address these processes, we ask the following questions:

- How do anti-freeze peptides self-assemble into nanotubes? What is the nanotube stability as a function of size?
• How do globular proteins associate into dimers? Which interactions play a role? What is the role of water along the association?

• Can we sample transitions with asymmetric barriers? Can we extract information about the transition state region?

1.7 Outline of the thesis

In this thesis we study the hydration layer dynamics and the association mechanisms of food and anti-freeze proteins. We do so by employing Molecular Dynamics and Transition Path Sampling. In chapter 2, we give a brief presentation of the MD and TPS methods. In chapter 3, we apply MD to address the dynamics of hydration water around native and misfolded α-lactalbumin. In chapter 4, we study the correlation between water structure and dynamics in the hydration layer of a type III ocean pout anti-freeze protein. In addition, in chapter 5, we investigate the stability and growth mechanism of self-assembling anti-freeze cyclic peptides. In chapter 6, we develop the spring shooting, a novel efficient transition path sampling move, and finally, in chapter 7 we elucidate the mechanism and role of solvent for β-lactoglobulin dimerization using Transition Path Sampling.
Bibliography


[38] S. Arrhenius, Z. Phys. Chem. 4, 96 (1889).


