A Molecular Perspective on the Dynamics and Entropy of Solvation Shell Water

A THESIS

SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS

OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

DEBASIS SAHA

20133260

DEPARTMENT OF CHEMISTRY
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE – 411 008

Dedicated to

My

Beloved Family…

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान, पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE
(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)
Dr. Homi Bhabha Road, Pashan, Pune, Maharashtra 411 008, India.

CERTIFICATE

Certified that, the work incorporated in the thesis entitled, "***A Molecular Perspective on the Dynamics and Entropy of Solvation Shell Water***" submitted by ***Debasis Saha*** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any deg ree or diploma from any other University or institution.

Date: 16th January 2017
Pune (MH), India

Dr. Arnab Mukherjee
(Thesis Supervisor)

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 14th December 2018 **Debasis Saha**

Pune (MH), India **Reg. No: 20133260**

ACKNOWLEDGMENTS

I acknowledge Indian Institute of Science Education and Research (IISER), Pune for providing high quality research environment and graduate fellowship. I am grateful to Prof. K. N. Ganesh, Former Director, IISER Pune and Prof. Jayant B. Udgaonkar, Current Director, IISER Pune for providing and maintaining excellent facilities for research. I would like to thank my thesis supervisor Dr. Arnab Mukherjee for the valuable guidance and mentoring during my Ph. D. I would specially thank him for the freedom he provided for independent thinking in the lab, which always helped me in my growth as a researcher.

I would also like to thank my Research Advisory Committee (RAC) members, Dr. Pankaj Mandal and Dr. Sarika Maitra Bhattacharyya, and also my former RAC member Dr. Suman Chakrabarty for their valuable suggestions and discussions during the annual meetings. Also, I would thank Dr. Pinaki Talukdar and Dr. Sujit K. Ghosh with whom I got the opportunity to work on new systems during my Ph. D.

I would like to give special thanks to my former and current lab members: Wilbee DS, Mandar Kulkarni, Reman Kr. Singh and Hridya VM for the various help, assistance and support during my Ph. D. journey. Also I would thank my other lab members: Abhijeet Gupta, Rahul Kumar, Amitosh Gautam and Atul Thakur and others from the department: Bappa Ghosh, Subrahmanyam Sappati, Avdhoot Datar for the various discussions and helpful interactions.

Finally, I would thank my family: my mother, father, brother and my sister-in-law for always supporting and believing in me. I especially thank my mother for her constant encouragement to be better in every aspect of life. I would like to give special thanks to my brother Asish Saha, for all the moral support and valuable suggestions at different stages of my Ph. D. life.

CONTENTS

List of Publications

1. **D Saha**, S Supekar, A Mukherjee*. Distribution of Residence Time of Water around DNA Base Pairs: Governing Factors and the Origin of Heterogeneity. *J. Phys. Chem. B* 2015, 119 (34), 11371-11381.

2. **D Saha** & A Mukherjee*. Impact of Ions on Individual Water Entropy. *J. Phys. Chem. B* 2016, 120, 7471-7479.

3. **D Saha***, M Kulkarni, A Mukherjee*. Water modulates the ultraslow dynamics of hydrated ionic liquids near CG rich DNA: consequences for DNA stability. *Phys. Chem. Chem. Phys.* 2016,18, 32107-32115.

4. **D Saha** & A Mukherjee*. Connecting diffusion and entropy of bulk water at the single particle level. *J. Chem. Sci.* 2017, 129,825-832.

5. **D Saha** & A Mukherjee*. Effect of Water and Ionic Liquids on Biomolecules. *Biophys. Rev.* 2018.

6. **D Saha** & A Mukherjee*. Water entropy in the supercooled state: Effect of dynamic arrest and local structures. (Manuscript under preparation).

7. T Saha, MS Hossain, **D Saha**, M Lahiri, P Talukdar*. Chloride-mediated apoptosis-inducing activity of bis (sulfonamide) anionophores. *J. Am. Chem. Soc.* 2016, 138 (24), 7558-7567.

8. S Mukherjee, A Kansara M., **D Saha**, R Gonnade, D Mullangi, B Manna, A. Mukherjee and S. K. Ghosh*. An Ultrahydrophobic Fluorous Metal-organic Framework Derived Recyclable Composite As A Promising Platform to Tackle Marine Oil Spills. *Chemistry A European Journal*, 2016.

9. A Roy, **D Saha**, A Mukherjee, P Talukdar.* One-Pot Synthesis and Transmembrane Chloride Transport Properties of C3-Symmetric Benzoxazine Urea. *Org. Lett.* 2016.

10. A Roy, **D Saha**, PS Mandal, A Mukherjee, P Talukdar*. pH-Gated Chloride Transport by Triazine-based Tripodal Semicage. *Chemistry A European Journal*, 2016.

11. S Sarkar, **D Saha**, S Banerjee, A Mukherjee*, P Mandal*. Broadband terahertz dielectric spectroscopy of alcohols. *Chem. Phys. Lett.* 2017, 678, 65-71.

ABSTRACT

The dynamic and thermodynamic properties of water remains an active area of research due to its significance in the biological processes. Although the pure water properties have been extensively studied through experimental and theoretical methods, water in the solvation shell represents a more realistic and complex scenario for naturally occurring water molecules. Using all-atom molecular dynamics simulations, the dynamics (diffusion coefficients and mean residence times) and thermodynamics of water molecules (entropy) in the solvation shell and bulk water have been investigated in my thesis. With the help of 360 different correlation functions, we showed that the key factor that governs the dynamics of water in a particular DNA groove is the position along the DNA. Extending this study to DNA in hydrated ionic liquid (IL) systems, we showed how the dynamics of water in these hydrated IL solutions causes significant changes in the dynamics of IL cations near different DNA base pairs. In thermodynamics, we investigated the entropy of individual water molecules around different cations and anions. After establishing the reliability of our method by showing the correlation with experimental solvation entropy, we discussed how cations and anions affect solvation shell water differently. We also show a significant contribution of entropy change in the solvation shell originates from the second solvation shell. Finally, we investigated the relation between dynamics and thermodynamics by studying water in the supercooled state. We could show that at the temperature where the dynamical transition takes place in water, the thermodynamic behavior of water (i.e., entropy) also changes. Further, our results provide new insights on the effect of structural polymorphism (high and low-density liquid) and fragile to strong crossover transition on water entropic behavior in its pure state. When the study is extended to the solvation shells of ions, the results correlate the structural polymorphism to the freezing behavior of water. Overall, the study aims to enrich the understanding of solvation shell water in addition to complementing previously known phenomena.

Chapter 1

Introduction

1. Uniqueness of Water

Water, known as the "matrix of life"[1] continues to uncover its unique properties in its pure form as well as in the solution phase. The qualities that make water so crucial compared to other such small molecules originate from its constituent atoms. Due to the large difference in their electronegativity, the oxygen and hydrogen atoms of water possess high charge density on themselves. This makes water exhibit qualities that makes it the prerequisite of life with most of the biological processes taking place in an aqueous environment.[2] Perhaps the most important property arising from the charge distribution on water molecules is a local interaction known as hydrogen bonds (H-bonds)[3]. The ability of water to form extended networks of fluctuating H-bond interactions with charged or polar moieties of biomolecules plays a significant role in the stabilisation and function of biomolecules[4-5]. In the bulk phase, while the maximum number of H-bonds a water molecule can form is four, water possesses slightly less than four H-bonds (~ 3.6) at ambient conditions. The other features of water that make it special compared to other liquids is its large dipole moment reflected in its high dielectric constant and the ability to form multiple structures of similar energies[6]. In addition to these, important processes such as molecular aggregation and protein folding in aqueous media is driven by an interaction known as hydrophobic effects[7], a phenomena yet to be understood completely. With all these properties, the study of different attributes of water in bulk phase and in the vicinity of various molecules has remained intriguing as well as challenging from a theoretical, simulation and experimental point of view.

Water behavior gets modified in the presence of different environments. To distinguish between the characteristics of water in the bulk phase and in the neighbourhood of solutes, the concept of hydration shell[8-9] has been suggested. The hydration shell includes the water molecules whose properties are significantly perturbed by the presence of a solute such as a biomolecule[10]. The number of affected water molecules and the extent of changes in the properties are governed by the nature of the solute. However, a quantitative picture of hydration shell properties transpires only through measurement of specific quantities. In a broad sense, the behavior of hydration shell water can be characterised by three main classes of properties: dynamic properties, thermodynamic properties and the structural properties. Although the

various quantities measured through experiments and simulations reveal different attributes of water, these properties are mostly interdependent on each other. Experimental measurements of dynamic properties such as translational or rotational diffusion, residence time, H-bond lifetime etc. require sensitive techniques that can distinguish the minor changes in the hydration shell water behavior compared to the bulk water. While recent advances in the field of experimental techniques have been able to probe the very small time and length scales involved in water dynamics, accurate interpretation of the observations form experiments remains a challenge. Molecular dynamics (MD) simulations in this regard has been proved to be advantageous where properties even at the individual molecule level are easily accessible. However, the MD simulation results are often dependent on the accuracy of the force fields used. Therefore a thorough investigation of hydration shell properties require a combination of both experimental and simulation studies. In this chapter of the thesis, various aspects of water behavior have been discussed with a focus on the existing knowledge in this field.

1.1 Dynamic Properties of Water

Depending on the interaction, the mobility of water molecules in the vicinity of a solute such as a biomolecule can either enhance or reduce. This creates differences in the dynamic properties of hydration water compared to the bulk water. The changes in the dynamic properties play crucial roles from a functional point of view of proteins and enzymes.[11] In the next section, the crucial aspects of hydration shell are discussed to emphasise its importance.

1.1.1 Significance of Water Dynamics

The majority of the cell volume remains occupied with water. The biomolecules perform their functions when they remain solvated in the aqueous medium. In fact, it has been found that for proteins to remain functional, a monolayer of water is required which quantitatively represents around 0.3 g of water per gram of protein.[12] Although there have been some exceptions found for this behavior,[13-14] several reports have claimed the dependence of protein dynamics and functions on the behavior of the hydration shell dynamics.[15-18] The concept of "slaving" has been termed in this regard to describe the coupled motions of solvent and proteins.[15] This notion has been established through several studies by observing the temperature dependence of water dynamics and protein fluctuations.[16-18] Young and co-workers have found the different

motions observed in proteins to be similar to the α and β-fluctuations in glasses.[16] In addition to proteins, the same has been observed in case of nucleic acids also in different studies[19-21] indicating the universal role of solvent dynamics for the functional aspects of biomolecules. The lack of structural transitions observed in case of biomolecules in their dry states for a given range of temperatures as shown in Fig. 1.1(A) gives a clear evidence of these facts.

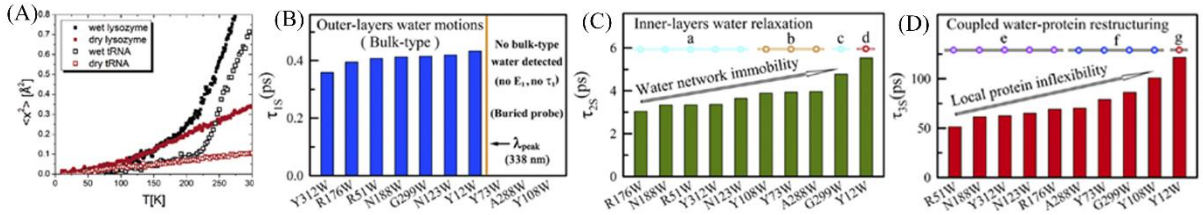

Figure 1.1 (A) The mean square fluctuation values for atoms of lysozyme and tRNA in their dry and hydrated states. While the hydrated molecules show a transition in the values at around 250 K, such transition is absent in the case of dry molecules. (B)-(D) The different solvent relaxation times observed in the study by Zhong and co-workers[22] around a protein. The three plots represent the ultrafast time scales (B), collective water/side chain reorientation timescale (C) and cooperative water/side chain restructuring timescale. The different factors responsible for the various timescales originate from partial charge (a), probe density (b), charge density (c), topological geometry (d and g) and intra-protein interactions (e and f). Reprinted with permission from ref. [19] and [22].

The evidence for the coupling of motions between water and biomolecules is not only limited to the structural transitions happening at a certain temperature region. Numerous reports have provided proofs for water mediated motions in proteins even at ambient conditions. Most notable of these are the studies by Zhong and co-workers[22-23] where the water dynamics and protein side chain motions have been investigated using a tryptophan scan with femtosecond spectroscopy. Different kinds of time scales have been measured through these studies: the first is in the order of hundreds of femtosecond corresponding to bulk like water motions, the second being the collective water/side chain reorientation with timescales of few picoseconds and the third being cooperative water/side chain restructuring in the timescales of tens of picoseconds. These timescales involving different amino acid residues are shown in Fig. 1.1(B)-(D). In other studies, using a combination of terahertz spectroscopy, X-ray absorption and MD simulation, Havenith and co-workers have shown that the changes in water dynamics and consequently the gradient of coupled protein-water motions play crucial role in enzyme-substrate interactions.[24] In addition to these effects, the hydration shell dynamics has been suggested to play roles in governing the antifreeze activity of some proteins. [25] These examples, along with several others

not discussed here, make the study of water dynamics essential in terms of both fundamental understanding as well as for controlling biological processes.

1.1.2 Measurement of Dynamic Properties of Water

A variety of quantities are used to describe dynamic properties of liquids at various conditions. While the measurement of self-diffusion coefficient (D) gives a direct measure of the mobility of water in the bulk phase or in the vicinity of biomolecules, several other dynamic quantities are measured through estimation of the timescales involved in the process. The value of D is related to the viscosity of the medium, η by the following relation given in Eq. 1.1, known as the Stokes-Einstein equation:

$$\mathrm{D} = \frac{\mathrm{k_B T}}{6\pi\eta \mathrm{r}} \qquad (1.1)$$

Where k_B is the Boltzmann's constant, T is the absolute temperature and r is the radius of radius of the spherical particle. However, value of D is calculated from either the mean square displacement (MSD) of atoms using the relation

$$D = \frac{1}{6t} \lim_{t\to\infty} \langle \| r_i(t) - r_i(0) \|^2 \rangle, \qquad (1.2)$$

or the value of D can be obtained from the velocity auto-correlation function (VACF) using the Green Kubo relation[26] given by,

$$\mathrm{D} = \frac{1}{\mathrm{d}} \int_0^\infty \langle \vec{\mathrm{V}}(0).\vec{\mathrm{V}}(\mathrm{t}) \rangle. \qquad (1.3)$$

Other ways to evaluate the dynamic properties of water is the measurement of mean residence time (MRT) of water molecules in the solvation shell of molecules. The MRT measures the time required for a water molecule to leave the solvation shell. A higher MRT value generally indicates stronger interaction for water molecule with the solute through electrostatic or H-bond interactions, while a lower MRT value indicates that water molecules diffuses easily from the solvation shell to the bulk environment.

While diffusion or residence time describes the translational motions of water molecules that are known to slave the protein motions, the other dynamic quantities through which water-network around solutes relaxes are its rotational motions. The reorientations of water molecules

are a result of such rotational motions which have been subject of numerous studies. The schematics of these motions are shown Fig. 1.2 and is adapted from ref. [27].

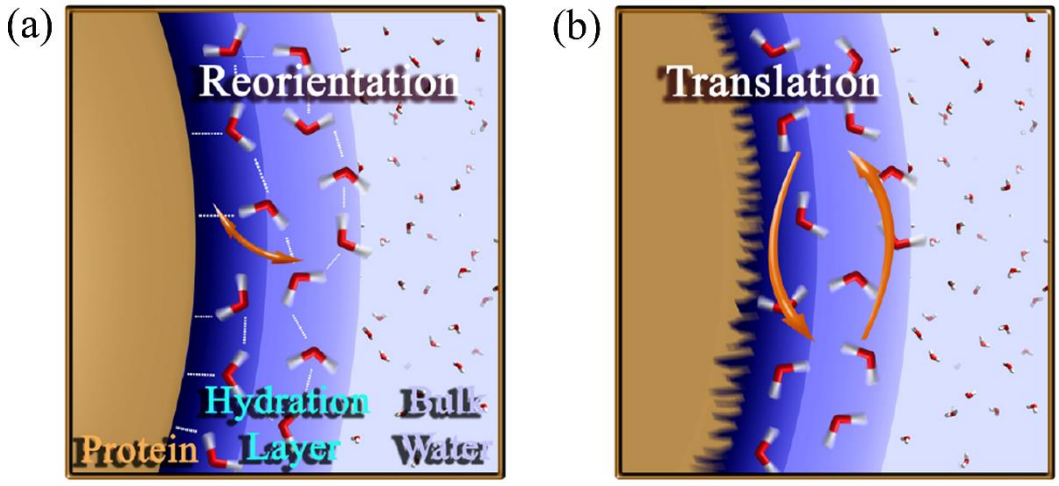

Figure 1.2 Schematic representation of (a) reorientation and (b) translational motion of water molecules in the hydration shell of a biomolecule such as protein. Reprinted with permission from ref. [27].

Similar to the translational diffusion of water molecules, the reorientation can be described using a diffusive pathway for water and the rotational diffusion coefficient can be evaluated from the measurement of angular mean square displacement $\langle \theta^2(t) \rangle$ using the following equation:[28]

$$D_R = \lim_{t\to\infty} \frac{1}{4t} \langle \theta^2(t) \rangle \qquad (1.4)$$

In general, the reorientation of water around different solutes is talked about in terms of the time required for a water molecule to reorient itself and it is known as the reorientation time of water. However the mechanism for this process have been debated. One of the description for water reorientation discusses it as an angular Brownian motion involving uncorrelated small angular steps. Such Brownian picture[29] often seems plausible since water remains in strong interaction environment with its neighbours. However, the observation of water reorientation reveals that the orientation of individual water O-H bonds show sudden large amplitude jumps rather than a gradual change. The orientation correlation functions within the diffusion model can be written as:

$$C_n(t) = \langle P_n[\mathbf{u}(0).\mathbf{u}(t)] \rangle, \qquad (1.5)$$

where P_n is the n-th-rank Legendre polynomial and $\mathbf{u}$ a vector attached to the molecule. The correlation function shows a monoexponential decay whose time constant is called the

reorientation time, τ_n. However, the value of τ_n obtained from D_R has often been inconsistent with the values calculated from MD simulations.[31]

A major improvement in the study of water reorientation came from the study by Laage and Hynes[30-31] which challenges many experimental data where water reorientation has been considered to follow a diffusive process. The mechanism by Laage and Hynes, known as jump-diffusion model, is shown in Fig. 1.3 where the breaking and formation of H-bonds have been proven to be a concurrent process rather than a sequence of small steps.

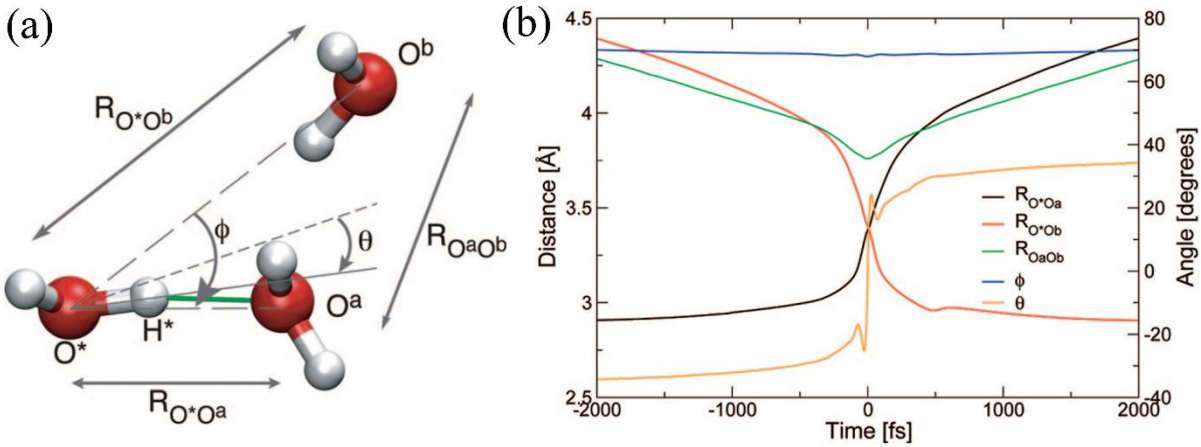

Figure 1.3 (a) The coordinates that describe the H-bond exchange event: $R_{O^*O^a}$, $R_{O^*O^a}$ and $R_{O^*O^a}$ represents the distances between the three water molecules, θ represents the angle between the O^*H^* bond and the bisector plane of $O^a\widehat{O^*O^b}$ and ϕ represents the $O^a\widehat{O^*O^b}$ angle. (b) The changes with time for these coordinates during the reorientation process. The time origin at t=0 is defined by θ=0. Reprinted with permission from ref. [31].

From the changes in the coordinates shown in Fig. 1.3(a), it can be seen in Fig. 1.3(b) how the breaking and formation of H-bonds crucial for water reorientation happens in a concerted mechanism rather than previously believed diffusive motions. This mechanism has been found to give a more accurate description for water reorientation motions around different solutes with variable timescales around them.

Other quantity that describes water dynamics around biomolecules is the H-bond lifetime of water. Two different types of time correlation functions (TCFs) are used in general to describe the timescale for H-bonds breaking: the continuous H-bond time correlation function, S(t), and the intermittent H-bond time correlation function, C(t)[32-33]. They are defined as the following:

$$S(t) = \frac{\langle h(t)H(t)\rangle}{\langle h\rangle} \qquad (1.6)$$

and

$$C(t) = \frac{\langle h(0)h(t)\rangle}{\langle h\rangle} \qquad (1.7)$$

Here $h(t)$ is a quantity that is unity when a particular pair of sites (such as water-water or an amino acid of a protein and a water) remain H-bonded, according to the H-bond definition criteria, and zero otherwise. The quantity $H(t)$ is unity when a tagged pair remain H-bonded from $t=0$ time to time t continuously and zero otherwise. The quantity $S(t)$ thus uses a stricter criteria for measurement of H-bond lifetime while C(t) gives the probability that a tagged H-bond remain intact at time t given that it was intact at time $t=0$. Thus C(t) does not takes into account whether the H-bond under investigation is broken at intermittent time or not. Once again, the timescale for the H-bond lifetime represents the strength of the H-bond between a water molecule and a solute such as an atom from a protein.

1.1.3 Experimental Investigations on Water Dynamics

Several experimental techniques are available that are capable of measuring the small timescales involved in water dynamics. However, often the difficulty with experimental techniques arises in separating the signal observed for water near a biomolecule from that of the bulk water. This makes the study of solutions with low concentrations of solutes very difficult. On the other hand, in case of higher concentration, the water dynamics is often influenced by the mixing of solvation shells of different molecules.

The nuclear magnetic resonance (NMR) technique have often been used to study water dynamics in the nanosecond timescale or slower. It can evaluate the water reorientation time by measuring the longitudinal spin relaxation rates of water oxygen or hydrogen isotopes, to which it is proportional.[34] The slow dynamics of water in the minor groove of DNA[35-36] and in the interior of protein[37] have been studied previously using this technique. The other NMR technique that has been widely used to study water dynamics is known as nuclear Overhauser effect (NOE). The results obtained from this, however, have often found mixed opinions.[38-39]

Incoherent elastic and quasi-elastic neutron scattering (EINS and QENS) are among the other techniques with high sensitivity required for the study of water in the solvation shell. However the results obtained from these are occasionally found to be ambiguous because of the accuracy of the model used to analysis.[40] The other aspect of neutron scattering experiment is that they

involve high concentration of solutes[41] or hydrated powders[42] for the measurement. Hence, the variation of concentration can lead to variable results. A recent study have used femtosecond pump-probe infrared spectroscopy to study the reorientation dynamics of water in protein solutions.[43] In spite the high sensitivity of this technique, resolving the dynamics for water present in the hydration shell is challenging due to its short picosecond lifetime.

While the above mentioned techniques are more specific for localized effect, several other techniques are available that probes collective motions of water molecules. Some of these techniques include depolarized light scattering (DLS), dielectric relaxation (DR), optical Kerr-effect (OKE) and terahertz (THz) spectroscopy. Although being successful at investigating the collective dynamic properties of water, some of the conclusions drawn from the studies involving these techniques are often subjected to debate. For example, the hydration shell around proteins at different concentrations has been claimed to extend up to 20 Å from the protein using THz spectroscopy.[44-45] However these conclusions obtained from these techniques often involve several approximations[46] with some other measurements suggesting otherwise.[47] Therefore care must be taken in carrying out and analysing the experiments for study of water dynamics at various environments.

1.1.4 Probing Water Dynamics through Simulations

Computer simulations have provided significant progress in understanding water behavior in pure state and in the solvation shells. The results obtained from experiments have also often been interpreted using either all-atom molecular dynamics simulation, Monte Carlo simulation, coarse grain simulations or *ab-initio* calculations. While computer simulations provide direct observations at very small time and length scales, the difficulty arises due to the increase in computational costs when the process under investigation involves larger system and long timescales.

The advantages of computer simulation methods for studying solvation shells come from the ease of study for water molecules confined between different biomolecules[48] or if the process involves very small timescales.[49] MD simulations also have been widely used to probe how fast or slow "biological water" is around different systems. While several studies have claimed conflicting results, [50-52] recent studies have found the water in the vicinity of biomolecules to be only 2-3 times slower than bulk water.[53] The anomalous nature of water around different solutes also makes the computer simulations important at an atomistic level. For example, the

dynamics of water is known to accelerate or decelerate around different ions compared to the bulk water. The mechanism for such behavior has been studied by Laage and co-workers[54] using all-atom MD simulations. However, simulations results are often dependent on the accuracy of the theory or the force field used in the study. An example of this can be realized from the study by Ding et al.[55] where the diffusion pattern in different salt solutions are only found to be consistent when ab initio MD simulations were carried out and not consistent with the simulation results obtained using the classical force fields.

While the accuracy of force fields used remain a source or error in the results obtained through simulations, the methods used to calculate certain property can also give misleading results in some cases. For example, a general method to calculate the mean residence time (MRT) of water around different solutes uses the survival probability function, $p_{\alpha,j}(t, t+t')$ which is one if a water j remains in the hydration shell of site α from time t' to $t'+t$, without getting out in this time interval and zero otherwise.[56] The average is taken over all the water molecules and over the whole timescale. The survival function was calculated in the following way:[56]

$$P_{\alpha}(t) = \sum_{j=1}^{N_{Water}} \sum_{t'} p_{\alpha,j}(t, t+t') \qquad (1.8)$$

A functional fit of this relation provides the MRT of water in the hydration shell. However, a water molecule can transiently leave the solvation shell and can return again before fully escaping the solvation shell. Since this is not considered in this method, the MRT values obtained can be underestimated in some cases. Therefore, care must be taken to obtain accurate values using such methods. In the studies discussed later in the thesis, an improvement of this method by Laage and Hynes[57] and the results obtained will be discussed in more details. While the above discussions are concerned with the dynamic aspects of water, the following topic will highlight the thermodynamic aspects of solvation shell water in various environments.

1.2 Thermodynamics of Solvation Shell Water

The reason why water is so important for biological systems comes largely from its unique interactions with different kinds of solutes. This factor plays a major role in governing the structures of biomolecules and in controlling various processes occurring in cellular environment. The occurrence of water in several crystal structures of biomolecules imply their strong interactions with the polar groups of these molecules. While these water molecules often mediate the interactions between different molecules such as proteins and DNA[48], often the

strongly bound water molecules alter the characteristics of the biomolecule. For example, the "spine of hydration" found in the crystal structure of DNA[58] have been known to interact strongly with the minor groove atoms. This has been argued to be the cause for higher free energy cost to bend DNA toward minor groove site.[59]

The other effect caused by water through its interaction with nonpolar moieties is known as hydrophobic effect. It is this interaction that causes the hydrophobic amino acids of a protein sequence to go to the interior of the protein structure, thus driving the protein folding process. Several protein-protein interactions are also governed by this effect where the hydrophobic patches of different proteins interact with each other and gets buried at the interface of the two proteins. One such example can be seen in Fig. 1.4(a) where the two-domain protein BphC enzyme (1dhy.pdb) has been shown in its complex form along with its interfaces.[60] The blue part in the bottom figure indicates the hydrophobic patch in the proteins which gets buried after the complex formation. The other important consequence of hydrophobic effect is the stability of cell membrane, as shown in Fig. 1.4(b). As seen from the figure, the hydrophobic alkyl chain remains hidden while the hydrophilic atoms are exposed to the aqueous environment.

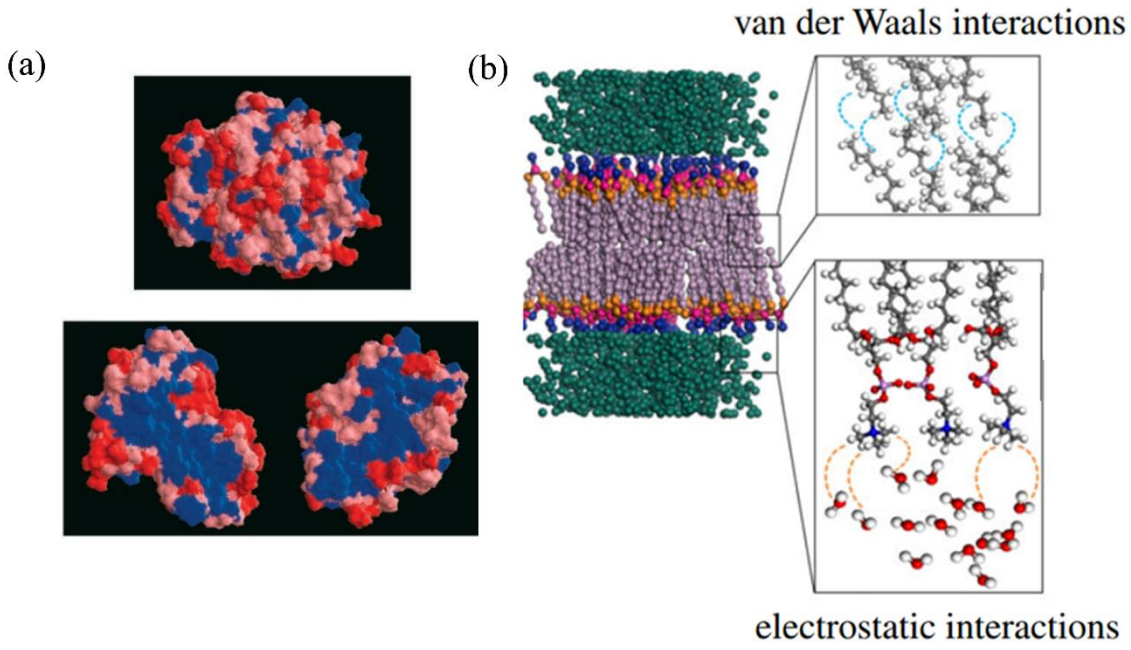

Figure 1.4 (a) The two-domain protein BphC enzyme in the upper panel and the separated complex with the hydrophobic patch at the interface shown in blue color in the bottom panel. Reprinted with permission from ref. [60]. (b) The structure of dipalmitoylphosphatidylcholine (DPPC) lipid bilayers with the hydrophobic and hydrophilic parts zoomed in along with their interactions. The green and grey color represent the water and alkyl chains while rest of the colors represent various atoms from the hydrophilic parts. Reprinted with permission from ref. [61].

While these effects contribute to the enthalpic component of the free energy for a process occurring in water, the other important effect of water comes from its entropy, which has been discussed in the next section.

1.2.1 Significance of Water Entropy

Several factors contribute to the molecular recognition processes taking place in biological systems. These include the enthalpic interactions between molecules[62] or the entropy of proteins originating from its conformational changes[63]. However, the solvent present both in the binding site and near the incoming molecule often plays crucial roles in modulating the free energy of the overall process. The entropy gained by the water molecules released in the binding process contributes favourably to the free energy of the process. This can be easily realised from the study by Hayashi *et al.*[64] where the binding of a ribonucleic acid aptamer with a partial peptide of a prion protein is shown to be driven by water entropy and is shown in Fig. 1.5. In this case, the energy decreases for the overall system since the binding of the two sites gets compensated by the increase in energy due to the loss in water-biomolecular interactions. As a consequence, the changes in the water thermodynamics plays a key role in the overall process.

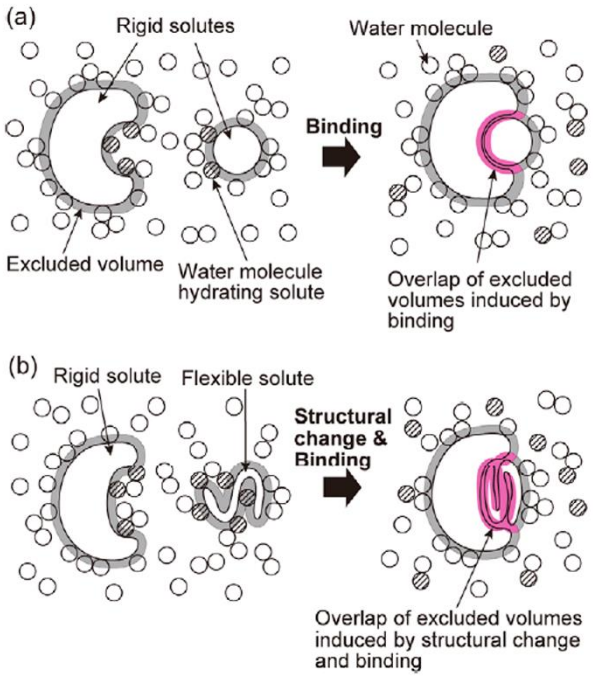

Figure 1.5 Schematic representation of the binding process of two solutes and the dehydrating process. The case shown in (a) has two rigid solutes and in the case shown in (b), one of the solute in flexible

and converts to a rigid form after the binding process. In both cases, the water molecules marked with shaded circles has gain in entropy. Reprinted with permission from ref. [64].

Such roles of water entropy in has been shown in several other studies on various protein systems[65-67] and near DNA[68]. Various attempts have been made over the years to predict the properties of water molecules near different solutes. While some of these methods are used to predict the water content in protein binding sites[69], other methods are used to calculate the thermodynamic properties of water at different environments. It has been shown recently that while water entropy, in general, is important in molecular recognition, it is a few specific water molecules whose entropy contribute significantly to the drug-protein binding[65] . Therefore, we are going to discuss on the specific topic of single water entropy in molecular recognition in the next subsection.

1.2.2 Role of Single Water Entropy in Molecular Recognition

The cavities of biomolecules often remain filled with water. In some cases, they remain strongly H-bonded to the solute. In general, the electrostatic interactions with the solute results in the stabilization of the water molecules inside a cavity. However, in the bound state both the translational and rotational motions of the water molecules remain restricted. This results in an unfavourable entropic state for the bound water molecules. An example for such a situation is shown in the study by Friesner and co-workers[70] in the outline hydration site of antibody DB3 where a water molecule remain H-bonded inside a hydrophobic cavity. This is shown in Fig. 1.6(a). For such a water molecule, release from the cavity will lead to enhancement in both translational and rotational entropy, leading to a favourable contribution toward free energy of binding for the incoming ligand. In cases where there are multiple water molecules present inside a cavity, the different water molecules are sometimes found to possess different value of entropy. This can be seen for the case of biotin binding cavity of streptavidin, where it has been shown that the five water molecules that form a pentameric arrangement have different individual entropy values[71] as shown in Fig. 1.6(b) and 1.6(c).

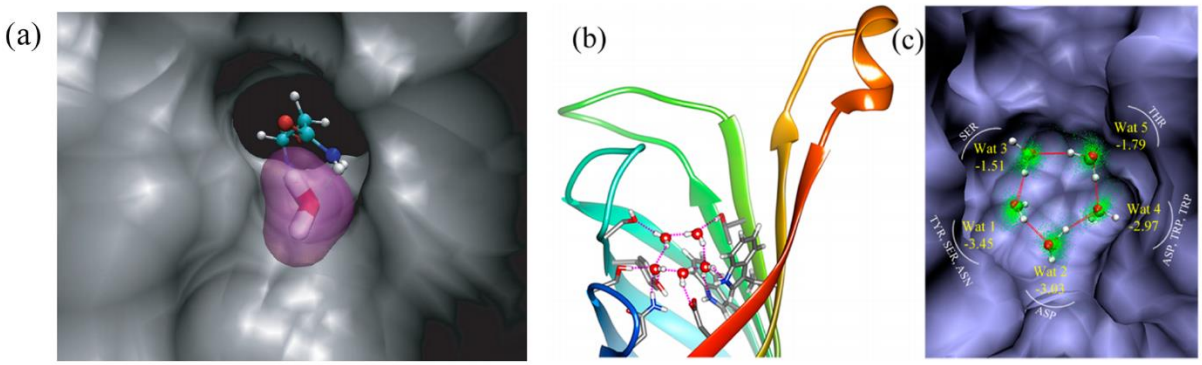

Figure 1.6 (a) The configuration of a water molecule in the hydration site of antibody DB3. The water is orientationally constrained through an H-bond with Asn-35 residue and is surrounded by hydrophobic residues. The purple shading shows the van der Waals radius for the water molecule. Reprinted with permission from ref. [70]. (b) The arrangement of five water molecules in the cavity of streptavidin in absence of biotin. (c) The total entropy values for the five water molecules shown in (b) in kcal/mol unit. The green dots represent the positions of water oxygen atoms in different frames and the protein cavity is shown with surface representation. Reprinted with permission from ref. [71].

The significance of such water molecules has been established in different studies where the entropy of water molecules has been used as a tool to design better ligand for a particular cavity. The study by Abel *et al.*[65] showed that for the active site of factor Xa, the specific design of ligand capable of removing particular water molecules from the active site can lead to favourable thermodynamics of overall ligand binding process. This is illustrated in Fig. 1.7 where the differently coloured spheres show the nature of different water molecules. It can be clearly seen from the figure that the cavity is occupied with several water molecules that can favour ligand binding entropically.

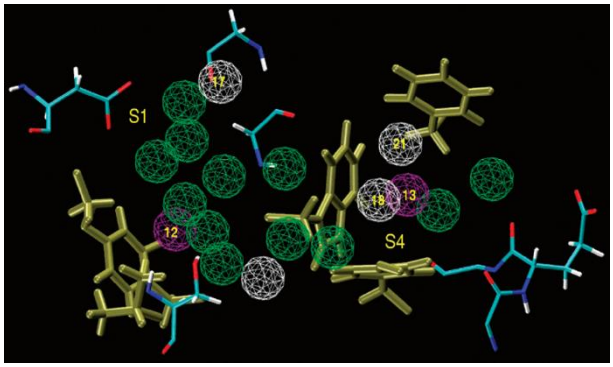

Figure 1.7 The water molecules in the active site of factor Xa in wireframe format. The S1 and S4 are the labels for the pockets shown in yellow and green, grey and purple colour represent water molecules which are expected to contribute entropically, energetically and both entropy and energetically respectively toward ligand binding. Reprinted with permission from ref. [65].

All these evidences show that exact knowledge of thermodynamics for individual water molecules in the vicinity of biomolecules is an important aspect in the development of better ligands. The various methods used for measurement of water thermodynamics has been discussed in the next section.

1.2.3 Methods Available for Water Entropy Calculation

A general estimate for the translational entropy of a monoatomic ideal gas can be obtained from Sackur-Tetrode (ST)[72] equation given as:

$$S^{id}_{Trans.} = C(T) + k_B \ln V, \qquad (1.9)$$

Where C is the momentum contribution and V is the configurational volume. Although the momentum contribution is unambiguous, the estimate for the configurational volume is often complex. For a real system such as water where the molecules interact with each other and its surrounding, the estimation of configurational volume is challenging and require an accurate method. An alternate method known as cratic entropy[73] uses an equation similar to ST where the configurational volume is measured from the information of concentration. However, the accuracy of this method has been a matter of debate.[74] Calculation of rotational entropy also faces similar challenges due the variation of water interaction with its environment. Several methods have been developed to estimate the translational and rotational entropy of water near solutes accurately, many of which has their root in the Inhomogeneous Solvation Theory (IST).[75-77] Using this approach, the thermodynamics of water, especially the entropic aspect of it, has been estimated in different ways in methods such as GIST,[78] STOW,[79] WaterMap,[70] and others.[80-81] In other methods, different statistical mechanical fundamentals are used (GCT,[82] 3D-RISM,[83] SZMAP,[84] 2PT[85] and WATsite[86]). Recently Kumar et al.[87] proposed tetrahedral entropy for water based on its structural characteristics. The details of these methods are not included in the discussion here. Although the estimates of thermodynamic properties obtained through these methods are shown to be mostly accurate, there are several drawbacks associated with these methods, some of which are discussed in the following section.

1.2.4 Drawbacks of Currently Available Methods

The disadvantages associated with the different available methods have been discussed in the study by Velez-Vega et al.[65] In several methods where the solvation shell thermodynamic properties are measured, the simulations are carried out with the fixed solutes[70, 75-77, 82] which may lead to incorrect solvation structures, and could possibly give inaccuracies in the measured values. In some cases, the highly occupied hydration sites are mostly studied[70, 80, 86] while the low-occupancy solvation and their thermodynamics are disregarded. Although the major contribution in the overall thermodynamics for the process happening in this kind of system may come from the highly hydrated regions, the contributions from other regions may also have significant contribution and may impact the ligand design and optimization. Often these difficulties are found to escalate for estimation of third and higher order corrections. With the various issues related to several of these methods, the search for newer methods with better accuracy and low computation cost remain an active area of research.

1.3 Structural Aspect of Water: Characteristics and Significance

Solvation shell water can exhibit different structures depending on its thermodynamic condition and solute properties. These structures often differ from each other in terms of energy and play crucial roles in the physical properties of water in its pure form and in the hydration shells. One of the most general characteristic for water structure is its H-bonds number. The general method for defining H-bond involves the donor and acceptor atoms to be within 3.5 Å of each other and the donor-Hydrogen-acceptor angle to be less than 30°. The number of such H-bonds depends on the environment of the water molecule under consideration. While a water can make maximum of four H-bonds, e.g., around ice at ambient condition, water makes on an average 3.6 H-bonds per water molecules.

The other commonly used structural parameter is known as the tetrahedral order parameter of water. The commonly occurring water is generally found to be in tetrahedral condition. The degree of tetrahedrality generally increases for water when the temperature is lowered. The local tetrahedral order parameter is used to quantify the tetrahedrality of water molecules, which is defined as:[88]

$$Q_k \equiv 1 - \frac{3}{8}\sum_i^3 \sum_{j=i+1}^4 [\cos\psi_{ikj} + \frac{1}{3}]^2, \qquad (1.10)$$

where ψ_{ikj} refers to the angle made by molecule k with its nearest neighbours i and j. The value of this parameter ranges between 0 and 1 corresponding to a complete non-tetrahedral structure to an entirely tetrahedral structures.

Other than these parameters, the general Steinhardt parameters of various orders are also used to characterise the structure of the solid form of water. The different aspects of water structures and their effect on the dynamics and thermodynamics of water at various conditions will be one of the focuses of the thesis.

1.4 Outline of the Thesis

In this thesis, the thermodynamics and dynamics of water in the solvation shell of different solutes have been investigated using computer simulations. The dynamic aspect of water has been investigated by probing the mean residence time (MRT) of water near DNA, while the thermodynamic aspect has been investigated through the calculation of entropy in the solvation shell. The relation between the dynamics and entropy of water have been investigated for water in its pure form and in the solvation shell in the supercooled state of water. A chapter-wise summary of the thesis is provided in the following:

Chapter 2 describes the method that has been employed in the thesis to study water dynamics and entropy.

Chapter 3 discusses the mean residence time (MRT) of water in the solvation shell of DNA base pairs to investigate the governing factors for water dynamics near DNA.

Chapter 4 describes the MRTs of ionic liquids near DNA base pair in hydrated ionic liquid systems to study how the chemistry of base pairs affect the dynamics of ionic liquids and water in the solution.

Chapter 5 shows the application of single water entropy calculation method developed within our group in the solvation shell of ions. The accuracy of the method has been shown in addition to obtaining new information on how cations and anions modulate water behavior in the solvation shell.

Chapter 6 describes the study of supercooled state of pure water through investigation of entropy and diffusion. Through analysis of structural aspects of water at different conditions,

the effect of structural polymorphism and dynamic transition on water entropy behavior has been discussed.

Chapter 7 shows how similar or dissimilar is the behavior of water in the solvation shell of ions in the supercooled state compared to water in its pure form. The entropy and structural aspects of water have been discussed in this context.

Chapter 8 summarises the main conclusions of the thesis. Also, few future directions have been discussed.

1.5 References

1. Szent-Gyorgyi, A. *In Cell-Associated Water*. Academic press: New York, 1979.
2. Rothschild, L. J.; Mancinelli, R. L. *Nature* **2001,** *409*, 1092.
3. Arunan, E.; Desiraju Gautam, R.; Klein Roger, A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary David, C.; Crabtree Robert, H.; Dannenberg Joseph, J.; Hobza, P.; Kjaergaard Henrik, G.; Legon Anthony, C.; Mennucci, B.; Nesbitt David, J., Defining the hydrogen bond: An account (IUPAC Technical Report). In *Pure and Applied Chemistry*, 2011; Vol. 83, p 1619.
4. Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006,** *35* (1), 389-415.
5. Ball, P. *Chem. Rev.* **2008,** *108* (1), 74-108.
6. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. Cambridge University Press: Cambridge, 2013.
7. Chandler, D. *Nature* **2005,** *437*, 640.
8. Kuntz, I. D.; Kauzmann, W. Hydration of Proteins and Polypeptides. In *Advances in Protein Chemistry*, Anfinsen, C. B.; Edsall, J. T.; Richards, F. M., Eds. Academic Press: 1974; Vol. 28, pp 239-345.
9. Bagchi, B. *Chem. Rev.* **2005,** *105* (9), 3197-3219.
10. Laage, D.; Elsaesser, T.; Hynes, J. T. *Chem. Rev.* **2017,** *117* (16), 10694-10725.
11. Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E. *Chem. Rev.* **2016,** *116* (13), 7673-7697.
12. Rupley, J. A.; Careri, G. Protein Hydration and Function. In *Advances in Protein Chemistry*, Anfinsen, C. B.; Richards, F. M.; Edsall, J. T.; Eisenberg, D. S., Eds. Academic Press: 1991; Vol. 41, pp 37-172.
13. Klibanov, A. M. *Nature* **2001,** *409*, 241.
14. Lind, P. A.; Daniel, R. M.; Monk, C.; Dunn, R. V. *Biochim. Biophys. Acta, Protein Struct.* **2004,** *1702* (1), 103-110.
15. Fenimore, P. W.; Frauenfelder, H.; McMahon, B. H.; Parak, F. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (25), 16047.
16. Fenimore, P. W.; Frauenfelder, H.; McMahon, B. H.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2004,** *101* (40), 14408.
17. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (13), 5129.
18. Kuo, Y.-H.; Chiang, Y.-W. *ACS Cent. Sci.* **2018,** *4* (5), 645-655.
19. Caliskan, G.; Briber, R. M.; Thirumalai, D.; Garcia-Sakai, V.; Woodson, S. A.; Sokolov, A. P. *J. Am. Chem. Soc.* **2006,** *128* (1), 32-33.

20. Kumar, P.; Yan, Z.; Xu, L.; Mazza, M. G.; Buldyrev, S. V.; Chen, S. H.; Sastry, S.; Stanley, H. E. *Phys. Rev. Lett.* **2006,** *97* (17), 177802.
21. Yoon, J.; Lin, J.-C.; Hyeon, C.; Thirumalai, D. *J. Phys. Chem. B* **2014,** *118* (28), 7910-7919.
22. Qin, Y.; Wang, L.; Zhong, D. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113* (30), 8424.
23. Qin, Y.; Jia, M.; Yang, J.; Wang, D.; Wang, L.; Xu, J.; Zhong, D. *J. Phys. Chem. Lett.* **2016,** *7* (20), 4171-4177.
24. Grossman, M.; Born, B.; Heyden, M.; Tworowski, D.; Fields, G. B.; Sagi, I.; Havenith, M. *Nat. Struct. Mol. Biol.* **2011,** *18*, 1102.
25. Meister, K.; Ebbinghaus, S.; Xu, Y.; Duman, J. G.; DeVries, A.; Gruebele, M.; Leitner, D. M.; Havenith, M. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (5), 1617.
26. Chandler, D. *Introduction to Modern Statistical Mechanics*. Oxford University Press: 1987.
27. Zhong, D.; Pal, S. K.; Zewail, A. H. *Chem. Phys. Lett.* **2011,** *503* (1), 1-11.
28. Mazza, M. G.; Giovambattista, N.; Starr, F. W.; Stanley, H. E. *Phys. Rev. Lett.* **2006,** *96* (5), 057803.
29. *J. Soc. Chem. Ind. London* **1929,** *48* (43), 1036-1037.
30. Laage, D.; Hynes, J. T. *Science* **2006,** *311* (5762), 832.
31. Laage, D.; Hynes, J. T. *J. Phys. Chem. B* **2008,** *112* (45), 14230-14242.
32. Stillinger, F. H. *Science* **1980,** *209* (4455), 451.
33. Rapaport, D. C. *Mol. Phys.* **1983,** *50* (5), 1151-1162.
34. Abragam, A. *The Principles of Nuclear Magnetism*. Oxford University Press: 1961.
35. Denisov, V. P.; Carlström, G.; Venu, K.; Halle, B. *J. Mol. Biol.* **1997,** *268* (1), 118-136.
36. Jóhannesson, H.; Halle, B. *J. Am. Chem. Soc.* **1998,** *120* (28), 6859-6870.
37. Kaieda, S.; Halle, B. *J. Phys. Chem. B* **2013,** *117* (47), 14676-14687.
38. Liepinsh, E.; Otting, G.; Wüthrich, K. *Nucleic Acids Res.* **1992,** *20* (24), 6549-6553.
39. Halle, B. *J. Chem. Phys.* **2003,** *119* (23), 12372-12385.
40. Qvist, J.; Schober, H.; Halle, B. *J. Chem. Phys.* **2011,** *134* (14), 144508.
41. Russo, D.; Murarka, R. K.; Copley, J. R. D.; Head-Gordon, T. *J. Phys. Chem. B* **2005,** *109* (26), 12966-12975.
42. Schirò, G.; Fichou, Y.; Gallat, F.-X.; Wood, K.; Gabel, F.; Moulin, M.; Härtlein, M.; Heyden, M.; Colletier, J.-P.; Orecchini, A.; Paciaroni, A.; Wuttke, J.; Tobias, D. J.; Weik, M. *Nat. Commun.* **2015,** *6*, 6490.
43. Brotzakis, Z. F.; Groot, C. C. M.; Brandeburgo, W. H.; Bakker, H. J.; Bolhuis, P. G. *J. Phys. Chem. B* **2016,** *120* (21), 4756-4766.
44. Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (52), 20749.
45. Born, B.; Kim, S. J.; Ebbinghaus, S.; Gruebele, M.; Havenith, M. *Faraday Discuss.* **2009,** *141* (0), 161-173.
46. Halle, B. *J. Phys. Chem. B* **2014,** *118* (36), 10806-10812.
47. Rosenfeld, D. E.; Schmuttenmaer, C. A. *J. Phys. Chem. B* **2011,** *115* (5), 1021-1031.
48. Chong, S.-H.; Ham, S. *J. Phys. Chem. Lett.* **2016,** *7* (19), 3967-3972.
49. van der Spoel, D.; van Maaren, P. J.; Larsson, P.; Tîmneanu, N. *J. Phys. Chem. B* **2006,** *110* (9), 4393-4398.
50. Li, T.; Hassanali, A. A.; Kao, Y.-T.; Zhong, D.; Singer, S. J. *J. Am. Chem. Soc.* **2007,** *129* (11), 3376-3382.
51. Golosov, A. A.; Karplus, M. *J. Phys. Chem. B* **2007,** *111* (6), 1482-1490.
52. Berg, M. A.; Coleman, R. S.; Murphy, C. J. *Phys. Chem. Chem. Phys.* **2008,** *10* (9), 1229-1242.
53. Furse, K. E.; Corcelli, S. A. *J. Am. Chem. Soc.* **2008,** *130* (39), 13103-13109.

54. Stirnemann, G.; Wernersson, E.; Jungwirth, P.; Laage, D. *J. Am. Chem. Soc.* **2013,** *135* (32), 11824-11831.
55. Ding, Y.; Hassanali, A. A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111* (9), 3310.
56. García, A. E.; Stiller, L. *J. Comput. Chem.* **1993,** *14* (11), 1396-1406.
57. Laage, D.; Hynes, J. T. *J. Phys. Chem. B* **2008,** *112* (26), 7697-7701.
58. Halle, B.; Denisov, V. P. *Biopolymers* **2000,** *48* (4), 210-233.
59. Ma, N.; van der Vaart, A. *J. Am. Chem. Soc.* **2016,** *138* (31), 9951-9958.
60. Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. *Science* **2004,** *305* (5690), 1605.
61. Youssefian, S.; Rahbar, N.; Lambert, C. R.; Van Dessel, S. *J. Royal Soc. Interface* **2017,** *14* (130).
62. Barratt, E.; Bingham, R. J.; Warner, D. J.; Laughton, C. A.; Phillips, S. E. V.; Homans, S. W. *J. Am. Chem. Soc.* **2005,** *127* (33), 11827-11834.
63. Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007,** *448*, 325.
64. Hayashi, T.; Oshima, H.; Mashima, T.; Nagata, T.; Katahira, M.; Kinoshita, M. *Nucleic Acids Res.* **2014,** *42* (11), 6861-6875.
65. Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008,** *130* (9), 2817-2831.
66. Beuming, T.; Farid, R.; Sherman, W. *Protein Sci.* **2009,** *18* (8), 1609-1619.
67. Singh, N.; Warshel, A. *Proteins Struct. Funct. Bioinf.* **2010,** *78* (7), 1724-1735.
68. Mukherjee, A. *J. Phys. Chem. Lett.* **2011,** *2* (24), 3021-3026.
69. Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2009,** *113* (40), 13337-13346.
70. Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (3), 808.
71. Sasikala, W. D.; Mukherjee, A. *J. Phys. Chem. B* **2014,** *118* (36), 10553-10564.
72. McQuarrie, D. A. *Statistical Mechanics* University Science Books: South Orange, NJ, 2000.
73. Irudayam, S. J.; Henchman, R. H. *J. Phys. Chem. B* **2009,** *113* (17), 5871-5884.
74. Holtzer, A. *Biopolymers* **1995,** *35* (6), 595-602.
75. Wallace, D. C. *J. Chem. Phys.* **1987,** *87* (4), 2282-2284.
76. Ashbaugh, H. S.; Paulaitis, M. E. *J. Phys. Chem.* **1996,** *100* (5), 1900-1913.
77. Lazaridis, T. *J. Phys. Chem. B* **1998,** *102* (18), 3531-3541.
78. Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. *J. Chem. Phys.* **2012,** *137* (4), 044101.
79. Li, Z.; Lazaridis, T. Computing the Thermodynamic Contributions of Interfacial Water. In *Computational Drug Discovery and Design*, Baron, R., Ed. Springer New York: New York, NY, 2012; pp 393-404.
80. Huggins, D. J. *Phys. Chem. Chem. Phys.* **2012,** *14* (43), 15106-15117.
81. Haider, K.; Huggins, D. J. *J. Chem. Inf. Model.* **2013,** *53* (10), 2571-2586.
82. Gerogiokas, G.; Calabro, G.; Henchman, R. H.; Southey, M. W. Y.; Law, R. J.; Michel, J. *J. Chem. Theory Comput.* **2014,** *10* (1), 35-48.
83. Sindhikara, D. J.; Hirata, F. *J. Phys. Chem. B* **2013,** *117* (22), 6718-6723.
84. Grant, J. A.; Pickup, B. T.; Nicholls, A. *J. Comput. Chem.* **2001,** *22* (6), 608-640.
85. Lin, S.-T.; Maiti, P. K.; Goddard, W. A. *J. Phys. Chem. B* **2010,** *114* (24), 8191-8198.
86. Hu, B.; Lill, M. A. *J. Comput. Chem.* **2014,** *35* (16), 1255-1260.
87. Kumar, P.; Buldyrev, S. V.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (52), 22130.
88. Errington, J. R.; Debenedetti, P. G. *Nature* **2001,** *409*, 318.

Chapter 2

Systems and Methodology

2.1. Systems

In this chapter, the systems on which the studies have been carried out in the thesis and the methods used has been discussed. Firstly the systems on which the investigations are done are discussed which includes DNA, ionic liquids, ions in water and supercooled state of water. This will be followed by the various methods used in the thesis.

2.1.1 Deoxyribonucleic acid (DNA)

DNA is known as the principal genetic information carrier for living organisms. The basic building block of DNA is known to be a nucleotide unit. It is composed of three major chemical fragments: a heteroaromatic nitrogenous base, a deoxyribose sugar molecule and a phosphate group. The four nitrogenous bases are categorized into purine bases (adenine (dA) and guanine (dG)) and pyrimidine bases (cytosine (dC) and thymine (dT) See Fig. 2.1). These mode of interaction between these bases are non-covalent in nature which includes hydrogen bonding and stacking interactions. Due to these interactions, the helical structure of DNA is formed.

In general, guanine forms three hydrogen bonds with cytosine and adenine forms two hydrogen bonds with thymine as shown in Fig. 1.1, thus forming rungs of the helical DNA ladder. The exocyclic carbonyl oxygen atoms and ring nitrogen atoms act as hydrogen bond acceptors and the exocyclic amine group and –NH ring group adjacent to carbonyl group acts as a hydrogen bond donor. These standard hydrogen bonding pattern of G:C and A:T base pairs are termed as Watson-Crick base pairs. The hydrogen bonding stabilization is close to -0.25 kcal/mol per base pair[1-4] and are cooperative in nature when multiple hydrogen bonds are considered.

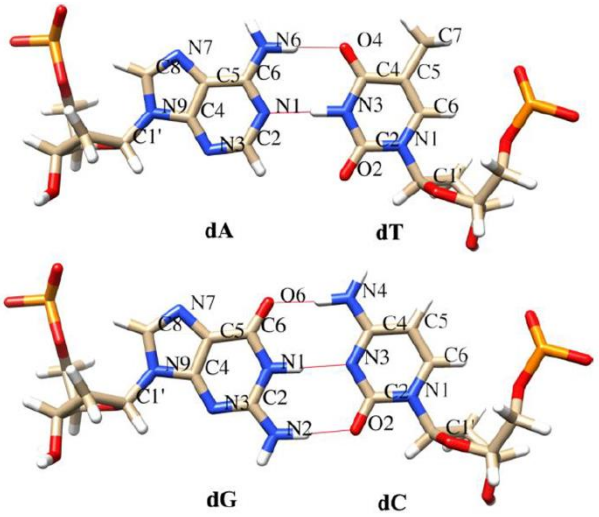

Figure 2.1 Figure showing specific atom types and Watson-Crick hydrogen bonding for (a) A:T base pair and (b) G:C base pair.

The aromatic base pairs have a tendency to stack over each other albeit in offset position leading to a conventional DNA helical geometry. The stacking arrangement of base pairs arises due to dispersion interactions between them as well as constraints imposed by sugar-phosphate backbone. The sugar-phosphate backbone holds the DNA framework stable. The flexible and hydrophilic sugar-phosphate backbone can potentially exhibit multiple conformations. For consecutive bases i-1, i, and i+1, these conformations can be observed using six backbone torsional angles O3'-[i-1]-P-O5'-C5' (α), P-O5'-C5'-C' (β) , O5'-C5'-C4'-C3' (γ), C5'-C4'-C3'-O3' (δ), C4'-C3'-O3'-P[i+1] (ε), C3'-O3'-P[i+1]-O5'[i+1] (ζ) and a dihedral angle χ around the glycosidic bond (N1/N9-C1'). For purine and pyrimidine bases, dihedral angle χ is defined as O4'-C1'-N9-C4 and O4'-C1'-N1-C2, respectively.

The glycosidic bond (N1/N9-C1') which connects a hydrophobic base to sugar plays an important role in defining conformational map of a certain nucleotide. The purine bases Adenine and Guanine can display both syn[5-6] ($50° \leq \chi \leq 80°$) and anti ($180° \leq \chi \leq 280°$) conformations. However, the pyrimidine bases Cytosine and Thymine exhibit only syn conformations due to steric clashes between base and sugar moiety. Thus, in a DNA helix, these various interactions such as hydrogen bonding, stacking interactions, steric repulsions and environmental parameters such as backbone-ions and backbone-water interactions eventually decide the fate of the DNA conformation.

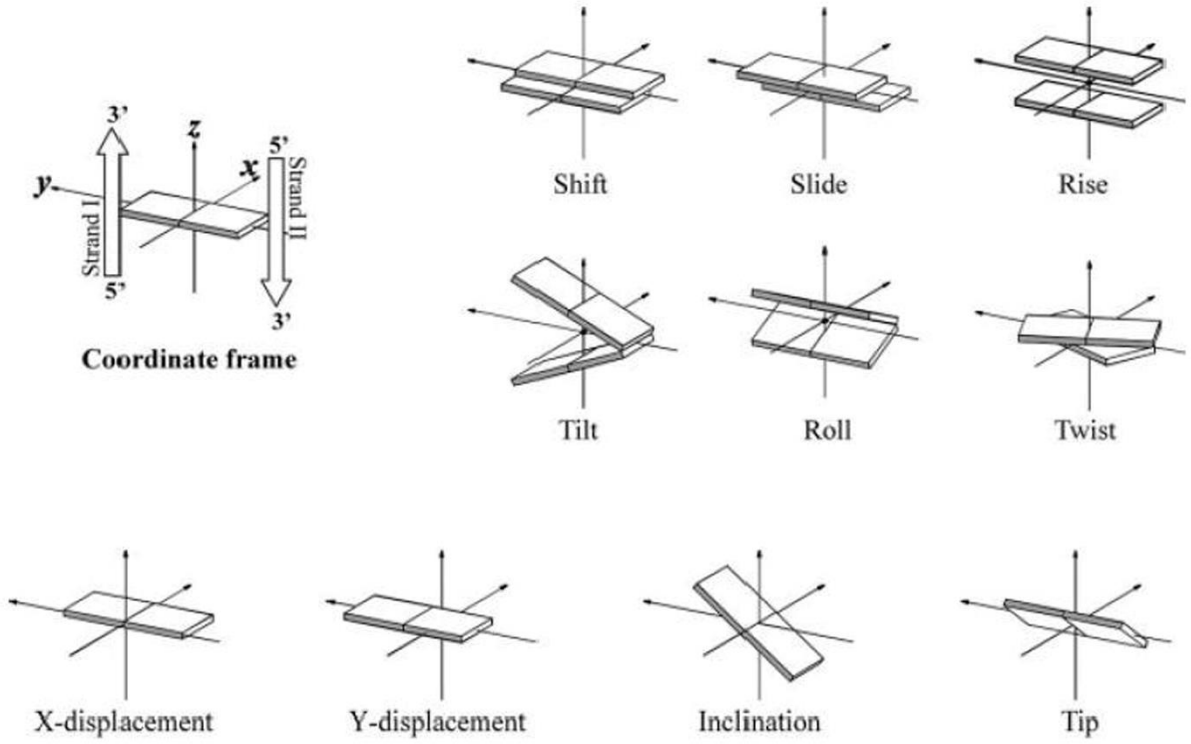

Figure 2.2 Figure showing the structural parameters of a dinucleotide step based on 3 translational motions (Shift, Slide and Rise) and 3 rotational motions (Tilt, Roll and Twist) according to block representation as proposed by Lu and Olson [7]. The figure is reprinted with permission from Ref. [8].

The bases Adenine-Thymine and Guanine-Cytosine interact via strong non-covalent hydrogen bonding interactions, thus forming a base pair. The base pair parameters such as X-displacement, Inclination display distinguishing values [8] for A-form (X-displacement = -4.81 Å, Inclination =19.8°) and B-form (X-displacement= 0 Å, Inclination = 0°). The DNA conformation depends on the various interactions present between consecutive base pairs and hence in the context of understanding DNA conformational map, a pair of base pair i.e. a dinucleotide step or a base-pair step represents the smallest possible unit. The chemical heterogeneity of each base step impose the minuscule variations in the structural properties of dinucleotide steps. These variations are recognizable using base-pair step parameters based on three translational parameters (Slide, Rise, Shift) and three rotational parameters (Roll, Twist, Tilt) as shown in Fig. 2.2. The knowledge of backbone torsional angles and sugar pucker along with these step parameters help to understand a complete conformational map of a dinucleotide step. Such meticulous information helps to anatomize a conformational diversity of whole DNA molecule into its corresponding individual dinucleotide steps, thus providing greater details to establish structure function relationship.

2.1.2 Ionic Liquids (ILs)

Ionic liquids (ILs) are salts that with high melting temperature. They are made up of large and unsymmetrical ions, because of which the packing between these ions are not proper. This causes these salts to remain in liquid form. One of the first report of room temperature ILs came in 1992 by Wilkes and Zaworotko[9]. Since then, recent years have seen several uses of ILs in various fields. One of the most important uses of ILs come in the field of biomolecular storage. Due to the low toxicity and less volatility of ILs, these salts serve as a suitable reagent for long term storage and applications of biomolecules. The other feature of ILs are their hygroscopic nature. Due to this, trace amount of water always remain present in IL solutions.

2.2 Methodology

2.2.1 All-atom Molecular Dynamics Simulation

In this thesis, all atom molecular dynamics (MD) simulations were used to study water at various environments. MD simulations use Newton's laws of motions to generate the velocity for atoms in the system. The forces are obtained from the potential energy originating from the interaction between atoms. MD simulations employ force fields to obtain the potentials, where the interatomic interactions are modelled via harmonic potential functions. The eq. 2.1 represents general form of force field,

$$U = \sum_{bonds} \frac{k_{bond}}{2}(r - r_{eq})^2 + \sum_{angles} \frac{k_{angle}}{2}(\theta - \theta_{eq})^2 + \sum_{torsions} \frac{V_n}{2}[1 + \cos(n\varphi - \gamma)] + \sum_{i>j}[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}], \quad (2.1)$$

Where k_{bond}, k_{angle} and V_n are force constants, n is multiplicity and γ is phase angle for dihedral angles. r_{eq} and θ_{eq} represent equilibrium values for bond and angle parameters. The last term of eq. 2.1 represents all non-bonded interactions: a combination of Lennard-Jones 6-12 potential and electrostatic interactions. The force field consists of constant values for k_{bond}, k_{angle}, V_n, n, γ, r_{eq}, θ_{eq}, A_{ij}, B_{ij}, q_i, q_j and ε whereas the variables r, θ, φ and R_{ij} are updated during simulation. From the trajectory obtained from the MD simulation, different analysis techniques were used to calculate various properties. The two main quantities that have been calculated in this study is the mean residence time of water and single water entropy. The methods used to obtain these quantities are discussed in the following section.

2.2.2 Mean Residence Time (MRT)

The mean residence time (MRT) measures the time taken for a water molecule to escape the solvation shell of a solute. Calculation of MRT involves monitoring the time a water molecule spends on an average near a site. So initially, at t = 0, probability that the molecule is at the site is 1. This probability, called as survival probability, decreases with time and finally goes to zero at some later time t.[10] A functional fit of the survival probability provides the time scale τ for the MRT. This approach, however, leads to an underestimation because a water molecule can transiently leave a site before it returns to it. Therefore, to allow such recrossing, Impey, Madden, and McDonald (IMM)[11] suggested a time lag of 2 ps for the water to re-join the site once it leaves. This allowed time-lag, however, is arbitrary, and MRT will depend on the lag.[12] Therefore, Laage and Hynes have used the stable state picture (SSP) of chemical reaction52 to calculate MRT of water around halide solution and showed that this approach is superior to IMM.[12] SSP approach requires the definition of a stable reactant and a stable product separated by a barrier. For such a two-state system, reactant and product regions are defined such that once the system enters the boundary of a region, it cannot return (i.e, forms a stable state). Such a definition takes care of transient escapes or recrossing events. Mathematically, the residence time via SSP approach may be given as the decay time of the complement of the probability that water is in stable reactant initially (t = 0) and forms a stable product at a later time t. The SSP procedure can be summarized as

$$C(t) = 1 - \langle p_R(0).p_P(t)\rangle \qquad (2.2)$$

Where p_R and p_P are the probabilities of being in the stable reactant and product states, respectively.[12] Absorbing boundary condition has been applied to the product state so that once the water reaches the product state, it is not considered further in the calculation. The method is schematically shown in Figure 2.3.

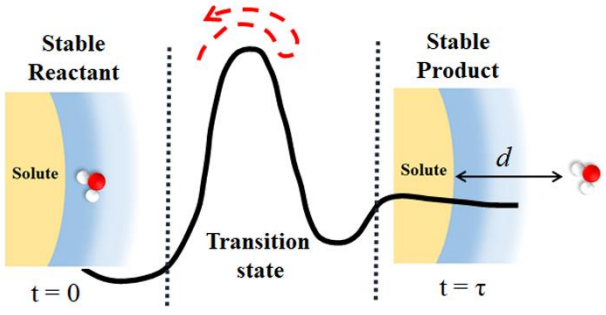

Figure 2.3 Schematic of the SSP approach used for the calculation of residence time. Vertical dashed lines show the stable state boundaries; reactant lies to the left of the left line, whereas the product is on the right of the right line.

This approach does not account for any state in between. Therefore, starting with the reactant, the approach allows any number of recrossing and any amount of time lag for recrossing before the product is formed. The criterion for defining the reactant and product region is discussed in chapter 3 of the thesis.

2.2.3 Single Water Entropy Calculation

In this thesis, the thermodynamic of water has been studied by calculating the individual water entropy values. A method developed within our group has been used for this purpose.[13] Our method for entropy calculations involves evaluation of translational ($TS_{Trans.}$) and rotational ($TS_{Rot.}$) entropy values for individual water molecule at a certain distance form the solute. However the main issue in this comes from the diffusive nature of water molecules, because of which a water molecule does not stay at its initial position throughout the whole trajectory. In our calculations, this problem is overcome by employing the method developed by Grubmüller and co-workers[14-15] known as permutation-reduction (PR) through which we obtain the effect of localized configuration space for an individual water molecule near a particular region. For the $TS_{Trans.}$ calculation, the permuted water molecules has been subjected to quasi-harmonic analyses. Therefore the covariance matrix of translational fluctuations in the x, y and z-direction has been obtained for the permuted water molecules. Diagonalization of this matrix provides three eigenvalues ($\lambda_i, i = 1, 2, 3$), which are then used to get the frequencies, $\omega_i = (k_B \mathrm{T}/\lambda_i)^{1/2}, i = 1,2,3$. Finally these frequencies are used in the expression for entropy of solid state quantum harmonic oscillator given in the following[16]:

$$S_{tr}^{QH} = k_B \sum_{i=1}^{3} \frac{\hbar\omega_i/k_B T}{e^{\hbar\omega_i/k_B T}-1} - \ln(1 - e^{\hbar\omega_i/k_B T}), \qquad (2.3)$$

Where $\hbar = h/2\pi$, h is Planck's constant, T is the temperature and k_B is the Boltzmann's constant. Each water molecules are permuted separately using our asymmetric permutation[17] approach. Considering the classical limit at high temperature ($\hbar\omega_i/k_B T < 1$), the above formula can be written in the following form:

$$S_{tr}^{QH} = k_B \sum_{i=1}^{3} 1 - \ln(1 - \hbar\omega_i/k_B T) = C(T) + k_B \ln V, \qquad (2.4)$$

where $C = (3 - \ln\left(\frac{4\pi}{3}\right) - \frac{3}{2}\ln\left(\frac{\hbar^2}{k_B T}\right)$ and $V = 4\pi/3(\frac{1}{2\sqrt{2}})(\lambda_1\lambda_2\lambda_3)^{\frac{1}{2}}$. Here $(\lambda_i)^{1/2}$represents the length of the principle axes of the ellipsoidal volume that a permuted water molecule occupies.

The anharmonic contribution arising from the interaction with solute has been taken in to account in the calculations by incorporating the method based on k^{th} nearest neighbour

distances, where k < (n-1) (n= sample points),[18] following Numata *et al.*[19] More details on this method can be found in ref. [13].

To calculate $TS_{Rot.}$ values, the angular distribution of permuted water molecules are used. The equation used to calculate $TS_{Rot.}$ values is given in Eq. 2.5:

$$S_{Rot.} = k_B \int p(\theta,\chi)c \ln\{p(\theta,\chi)c\} \sin\theta \, d\theta \, d\chi, \qquad (2.5)$$

Where $p(\theta,\chi)$ is angular distribution for a permuted water molecule and c is a normalization factor; θ represents the angle between water dipole vector and the vector between water oxygen and the solute; χ represents the angle between H-H vector of the water and the normal to the plane defined by the dipole vector and the solute-oxygen vector. These angles are shown in Fig. 2.4. The reason for not using quasi-harmonic entropy calculation method for rotational motion is that the moment of inertia associated with angular motion of water molecule relative to a solute is ambiguous. Also, the use of a harmonic approximation is inappropriate for a freely rotating molecule.[20]

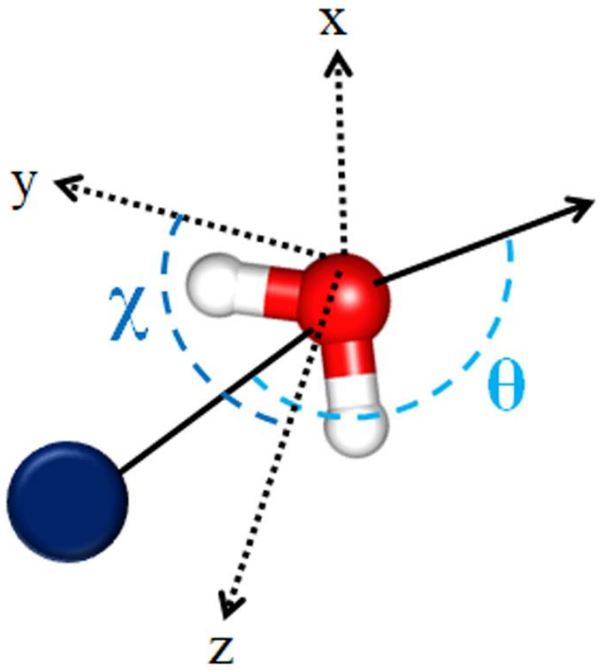

Figure 2.4. Schematic diagram showing the two angles θ and χ used for the calculation of rotational entropy. θ represents the angle between the dipole vector of water and the vector from water's oxygen to the solute shown as blue sphere. χ represents the angle between the H−H vector of water and the normal to the plane defined by the dipole vector of water and the solute-oxygen vector.

2.3 References

1. Freier, S. M.; Sugimoto, N.; Sinclair, A.; Alkema, D.; Neilson, T.; Kierzek, R.; Caruthers, M. H.; Turner, D. H. *Biochemistry* **1986,** *25* (11), 3214-3219.
2. SantaLucia, J.; Kierzek, R.; Turner, D. H. *Science* **1992,** *256* (5054), 217.
3. Serra, M. J.; Turner, D. H. Predicting thermodynamic properties of RNA. In *Methods in Enzymology*, Academic Press: 1995; Vol. 259, pp 242-261.
4. Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001,** *30* (1), 1-22.
5. Sychrovsky, V.; Foldynova-Trantirkova, S.; Spackova, N.; Robeyns, K.; Van Meervelt, L.; Blankenfeldt, W.; Vokacova, Z.; Sponer, J.; Trantirek, L. *Nucleic Acids Res.* **2009,** *37* (21), 7321-7331.
6. Karplus, P. A.; Shapovalov, M. V.; Dunbrack, R. L.; Berkholz, D. S. *Acta Crystallogr D Biol Crystallogr* **2008,** *64* (Pt 3), 335-336.
7. Lu, X.-J.; Olson, W. K. *Nucleic Acids Res.* **2003,** *31* (17), 5108-5121.
8. Lescrinier, E.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.* **2003,** *31* (12), 2975-2989.
9. Wilkes, J. S.; Zaworotko, M. J. *J. Chem. Soc., Chem. Commun* **1992,** (13), 965-967.
10. García, A. E.; Stiller, L. *J. Comput. Chem.* **1993,** *14* (11), 1396-1406.
11. Impey, R. W.; Madden, P. A.; McDonald, I. R. *J. Phys. Chem.* **1983,** *87* (25), 5071-5083.
12. Laage, D.; Hynes, J. T. *J. Phys. Chem. B* **2008,** *112* (26), 7697-7701.
13. Sasikala, W. D.; Mukherjee, A. *J. Phys. Chem. B* **2014,** *118* (36), 10553-10564.
14. Reinhard, F.; Grubmüller, H. *J. Chem. Phys.* **2007,** *126* (1), 014102.
15. Reinhard, F.; Lange, O. F.; Hub, J. S.; Haas, J.; Grubmüller, H. *Comput. Phys. Commun.* **2009,** *180* (3), 455-458.
16. Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001,** *115* (14), 6289-6292.
17. Mukherjee, A. *J. Phys. Chem. Lett.* **2011,** *2* (24), 3021-3026.
18. Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. *Structure* **1995,** *3* (6), 581-590.
19. Li, Z.; Lazaridis, T. *J. Phys. Chem. B* **2005,** *109* (1), 662-670.
20. Carlsson, J.; Åqvist, J. *J. Phys. Chem. B* **2005,** *109* (13), 6448-6456.

Chapter 3

Distribution of Residence Time of Water around DNA Base pairs: The Origin of Heterogeneity

3.1 Introduction

Understanding the dynamics of water near DNA is essential for the role these water molecules play in DNA replication and repair,[1] molecular recognition by DNA[2] and protein-DNA interaction.[3] One of the routes of dynamical relaxations of biological water near DNA is through an exchange with the bulk. This can be probed through mean residence time (MRT), which measures how long water molecules stay near a biomolecule. However, the definition of MRT may vary. MRT is interpreted as a measure of the delay between successive translational hops[4] or as the second-order orientational relaxation.[5] The range of MRT can have a large variation for different systems. First evidence of a long MRT, 10 ns to 10 ms for protein-interior water and 300−500 ps for surface-bound water, came from a combined nuclear Overhauser effect (NOE) and its rotating frame equivalent (ROE).[6] However, subsequent 2H and 17O nuclear magnetic relaxation dispersion (NMRD) studies[7] and 1H spin−lattice relaxation[8] indicated a much smaller value (30−50 ps). The reason for this discrepancy was attributed to the contribution from very slow and buried water molecules in protein to the average water time scale.

In the present study, we have focused on MRT of water in the solvation shell of DNA, or more specifically around different dinucleotide steps of DNA. DNA is an important biomolecule responsible for the growth and survival of life. The physiologically common B-form has a narrow and deep minor groove, wide and shallow major groove, and charged phosphate groups in a perfectly ordered arrangement. These structural features, which depend both on the DNA sequence and surrounding water, determine distinctly different structural and dynamical behavior of water molecules that gave rise to variety of opinions recently on DNA solvation[9-14].

The heterogeneity in water behavior can be easily realized around DNA in spite of it having a more regular structure compared to proteins. The residence time reported by different groups in the grooves of DNA has also been different. Liepinsh et al., using NOE/ROE, reported a much longer MRT of water (>1 ns) in the minor groove compared to that in the major groove (<500 ps) at lower temperature 10 and 4 °C.[15] In another study, Jacobson et al.

reported residence times >0.5 ns for hydration water, which were detected by positive water−DNA cross peaks in NOESY,[16] and the absence of long residence time (>1 ns) has been claimed by Sunnerhagen et al. around trp operator DNA.[17] At room temperature (27 °C), however, MRT of water in the minor groove of Dickerson dodecamer (DD) was found to be 200 ps, while it was estimated to be 140 ps in the major groove.[18] NOE and NMRD studies carried out for different DNA sequences in the presence and absence of minor groove binding drug netropsin[19] indicated that mostly long-lived water molecules come from the "spine of hydration" existing in the minor groove, and therefore, groove width was argued to play a role.[20] Generally, the time scales are found to be much longer for MRT than the properties probed by TRSS, dielectric relaxation, and so forth. However, these studies give an average behavior of water and do not provide any site-specific information.[13] Moreover, they are limited to study only specific DNA sequences as water around the minor groove of GC base pairs gives no NOE, indicating a shorter residence time.[21]

Computational studies can probe directly into the behavior of water around a particular site. However, the previous computational studies of MRT have primarily focused on proteins,[22-26] and rarely on DNA.[27] Moreover, most of the studies on water dynamics around DNA and proteins were focused on specific systems. Therefore, they do not provide a general picture about the origin of water dynamical behavior. Recently, two groups[28-29] calculated orientational relaxation around various sites of proteins and thus captured a distribution of possible relaxation time scale indicating an inherent heterogeneity in the distribution of dynamical properties as well as energetics of water across various sites of proteins. [28-29] Such behavior may be the consequence of both the topology[30-31] and the chemistry (i.e., electrostatic nature)[32-33] of the hydration site.

In a DNA system, both the topography (groove width/depth) and the chemistry of nucleobases and their sequential arrangement may vary. Therefore, it is quite non-intuitive how the behavior of water locally around a particular DNA dinucleotide will manifest.[34-35] This study thus aims at a general understanding of the water residence time around DNA base pairs. We have tried to address the question as to what governs MRT around DNA: the chemistry of the base pairs or the topography of the DNA. Therefore, we have calculated MRT, following the prescription by Laage and Hynes[36] based on the stable state picture (SSP) formulation[37] of Hynes and co-workers, around several DNA dinucleotide steps located at various positions along DNA. We have standardized our method for the calculation of MRT on the DD system. Subsequently, we have applied the method to several other DNA systems to find the governing factors of MRT. Apart from the chemistry and topography, this study also finds a new universal

correlation with respect to the position of site along the DNA where water molecules near the base pairs in the middle of the DNA have higher MRT that decreases continuously toward the terminals. We propose a simple kinetic model that captures not only the time scales of water escape from DNA but also the origin of this observed correlation.

3.2 Computation Details

3.2.1 Sequence Generation

To carry out the study, we have tested our method first on the Dickerson dodecamer (DD)[38] system with pdb id 1BNA. To further investigate the origin of heterogeneity in the MRT values, 20 DNA sequences were used, each having 12 base pairs. The sequences for these 20 DNAs are given in the following:

1. CGTGGAGATACG,
2. CGATATACGGCG,
3. CGGCATGAACGC,
4. CGATTACCTGCG,
5. CGATAGAACGCG,
6. CGCAAGCACGCG,
7. CGGCTGCTACGC,
8. CGGTATGCTACG,
9. CGGCCGAAAACG,
10. CGATAGCCGACG,
11. GCACTATGAAGC,
12. GCGCCGCCGTGC,
13. GCGCGTATACGC,
14. GCTAATCGGCGC,
15. GCATTCTTGGGC,
16. CGGTTAGATGCG,
17. GCGCATAGTAGC,
18. GCCTGATTAGGC,
19. CGCCAAAACTCG,
20. GCGAATCTGGGC.

The DNA sequences were chosen in such a way that each base pairs remain in equal numbers at various positions. These DNA sequences were processed using the NAB program[39] to generate their structure in B-DNA form. Xleap in AMBERTools[39] was used with AMBER99/parmbsc0 force field[40-41] to generate the topology and coordinate files. Subsequently, the coordinates and topologies obtained were converted to GROMACS[42] format using the amb2gmx.pl program.[43]

3.2.2 Simulation Details

All the DNA sequences were solvated using TIP3P[44] water molecules. The box dimensions were chosen such that there was a 10 Å thick solvation shell around each DNA system. The system was neutralized by replacing 22 water molecules by 22 sodium ions. These DNA systems were energy minimized using the steepest descent method followed by heating up to 300 K using a Berendsen thermostat48 with a coupling constant of 0.2 ps and a harmonic restraint of 2.5 kcal/mol.$Å^{-2}$ on heavy atoms of DNA. This was followed by a gradual sequential reduction of restraint to 50 cal/mol.$Å^{-2}$ in seven simulations, 50 ps each at constant temperature 300 K and constant 1 bar pressure using a Berendsen thermostat and a barostat,[45] respectively. The final equilibration without any position restraints was carried out for 1 ns at constant temperature 300 K and constant pressure 1 bar using Nose−Hoover thermostat[46-47] and Parrinello−Rahman barostat,[48] respectively, with a coupling constant of 0.2 ps for each. Electrostatic interactions were treated using PME electrostatics[49] with 10 Å cut-off. The van der Waals cut-off was set to 10 Å. A final molecular dynamics run of 10 ns was carried out for all the 20 DNA systems and DD system with similar treatment of temperature, pressure, electrostatics, and vdW as in final equilibration. All the molecular dynamics treatment was done using GROMACS package[42] with constrained bonds and time step of 2 fs.

3.2.3 Definition of Stable States

As discussed in chapter 2, the calculation of MRT values require defining the stable reactant and product states. Since in this case, the focus is on the effect of chemistry and topography of DNA base pair steps, a total of 360 sites (a particular dinucleotide step in the minor or major groove) are identified around which MRT calculations were carried out. Also calculation of MRT values were carried out near DD system. When a water molecule is close to a particular site, it is defined as reactant. For the product, we set the criteria that the water molecule must be far away not only from the site but also from the whole DNA. The reason for such a definition is that a water molecule in principle can move to another base pair and not to bulk after leaving the site, and in turn, it will contribute positively toward MRT in the experiment.

The precise definition of stable state boundaries requires the knowledge of the hydration layer in different grooves. Typically radial distribution function provides a good measure of the solvation shell. However, because DNA is not spherical, we used proximal radial distribution function (PRDF)[50] of water molecules around a particular site. PRDFs measure water distribution perpendicular to the local surface of the site.[51] For our purpose, we

have selected the atoms in the major and minor grooves of DNA and calculated the PRDFs by utilizing a faster grid-based algorithm proposed by Makarov et al.[52] Figure 3.1 shows some representative PRDFs of water molecules around different dinucleotide steps in the major (Figure 3.1a) and minor grooves (Figure 3.1b) of DD. The deeper and narrower character of the minor groove results in a narrower and more pronounced first solvation shell, along with a broad second solvation shell. Distribution of water on the major groove has a smaller difference in the structure of the two solvation shell; both are less intense than the first solvation shell of the minor groove. The PRDF plots show that the distribution of water is different grossly between major and minor grooves. The difference is, however, less among different base pairs in the same group, especially in the position of the peaks and troughs of the solvation shells. From the PRDF, we can obtain the underlying free energy or potential of mean force (PMF) using the relation[53]

$$\mathrm{PMF}(r) = -k_B T \ln \mathrm{PRDF}(r), \qquad (3.1)$$

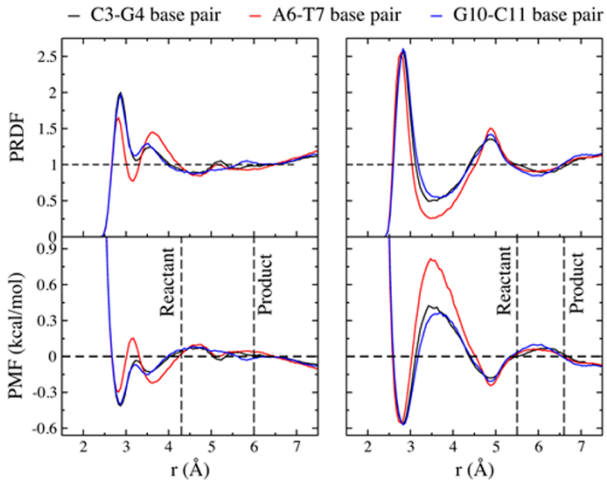

Figure 3.1 Proximal radial distribution functions (PRDF) of water around different base pairs of Dickerson dodecamer in the (a) major and (b) minor grooves are shown in the top panel. Corresponding potential of mean forces (PMFs) for major (c) and minor grooves (d) are shown in the bottom panel.

Figure 3.1c and 3.1d show the PMF of water around different sites of major and minor grooves, respectively. Based on the PMF, escape of water from base pair to bulk can be thought to follow the following kinetic behavior

$$S1 \rightleftharpoons S2 \rightarrow \mathrm{Bulk} \qquad (3.2)$$

Where S1 and S2 denote the first and second solvation shells, respectively. From the PMF, we can now define the reactant and product stable states. If d_{min} is the minimum distance of a water molecule from a site of the DNA, then the reactant region is defined where $d_{min} < 5.5$ Å and product region is defined when $d_{min} > 6.6$ Å for a site in the minor groove. For the major groove

sites, the reactant region is defined where d_{min} < 4.3 Å, and the product region is defined where d_{min} > 6 Å. In both cases, water molecules in the first and second solvation shell are considered to be within the reactant region. Figure 3.1c,d show the stable state boundaries for the major and minor groove sites, respectively. Note that these boundaries are defined between the barrier and stable states of the PMF, as required by the SSP approach.[36] Because the peak positions match between different base pair steps, we used the same condition across all DNA systems. Choice of the cut-offs plays an important role in the residence time calculation. Similar cut-off used in the major and minor groove was shown earlier to produce similar residence time in both grooves.[17]

3.2.4 Local and Directional Velocity Auto-Correlation Function (VACF)

For the kinetic model, discussed later, we used diffusion coefficients along the DNA and perpendicular to the DNA in the first and second solvation shell to calculate the rate constants. Diffusion coefficients were calculated from the velocity autocorrelation function using the following general relation[54]

$$D = \frac{1}{d}\int_0^\infty \langle \vec{V}(0).\vec{V}(t)\rangle, \qquad (3.3)$$

Where d is the dimensionality, D is the diffusion coefficient, and $\vec{V}$ is the velocity vector. For our study, we needed to calculate diffusion coefficient of water at a particular region along a particular direction. Therefore, we first created a local body fixed triad axes system in DNA, similar to what we recently used to study localized B → A DNA transition,[55] where $\hat{x}$ axis points to the minor groove and $\hat{z}$ axis aligns roughly along the DNA helical axis. The vectors are shown in Fig. 3.2.

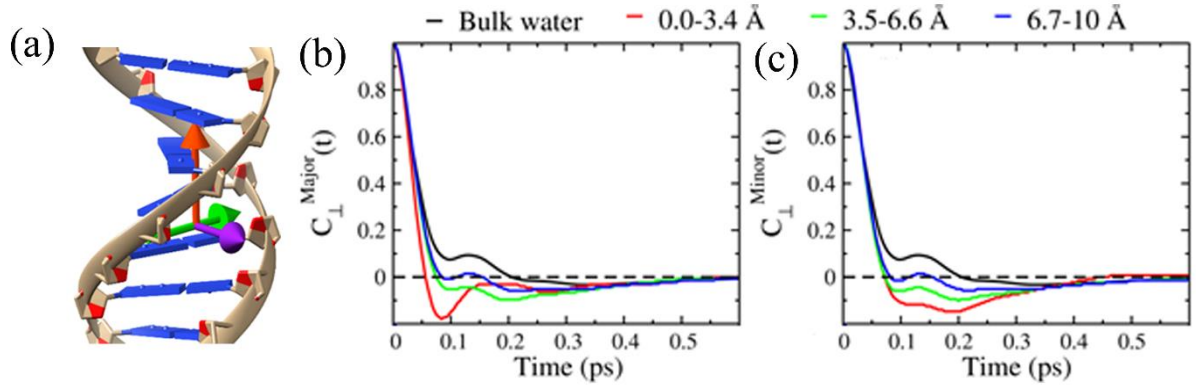

Figure 3.2 (a) The three vectors constructed on DNA to calculate VACF along different directions. The vector along perpendicular direction of DNA, denoted here as $\hat{x}$, is constructed from the center of mass (COM) of four bases to the COM of two 3' sugar groups. A vector along parallel to base pair, denoted as $\hat{y}$ here, is constructed from the COM of two C' atoms of sense strand to the two C' atoms of the

antisense strand. The cross product of $\hat{x}$ and $\hat{y}$, denoted as $\hat{z}$, is the vector parallel to DNA axis. $\hat{x}$, $\hat{y}$ and $\hat{z}$ are shown with violet, green, and red color. The VACF plots at various distances in the perpendicular direction of DNA for the major (b) and minor (c) groove.

Following the PRDF, we defined the first solvation shell water to be at a minimum distance within 3.4 Å; for the second solvation shell, water molecules at a minimum distance from the DNA within 3.4 to 6.6 Å were considered for the minor groove. The velocity of the water molecules that reside continuously for 2 ps within a given distance limit were projected onto $\hat{z}$ and $\hat{x}$ axes to calculate the VACF and in turn diffusion coefficient parallel and perpendicular to DNA, respectively. The VACF plots in the perpendicular direction of DNA are shown in Fig. 3.2(b) and 3.2(c) for major and minor groove respectively.

3.3 Results and Discussions

We standardized the SSP approach and the stable state cut-offs on the DD system first because DD is a well-studied system.[21] It also contains a narrow minor groove[56] and therefore expected to provide an upper-bound to the MRT. Subsequently, we applied the method to other 20 different DNA systems to obtain a distribution of residence time around different base pairs.

3.3.1 Calculation of Residence Time around Dickerson Dodecamer (DD)

Although our goal is to calculate water MRT around base pair steps (two base pairs) only, we first analysed the effect of the size and sequence variation of the sites. To do that, we used the SSP approach to calculate the correlation function C(t) using Eq 2.2. C(t) was fitted generally to a biexponential function of the following form after removing the initial fast decay (10 ps)

$$C(t) = C_1 e^{-t/\tau_1} + C_2 e^{-t/\tau_2}, \qquad (3.4)$$

Where C1 and C2 are the amplitudes of time constants τ_1 and τ_2, respectively. τ_1 is taken to be the faster, and τ_2 is taken to be the slower time scale by convention. Figure 3.3 shows a typical correlation function and biexponential fit for a site in the major and minor grooves of DD. The overall MRT is obtained by weighted average of both time scales as

$$\tau = \frac{C_1\tau_1 + C_2\tau_2}{C_1 + C_2}, \qquad (3.5)$$

Table 3.1 shows the individual time constants and the overall MRT of water around various sites in the major and minor groove of DD. Note that most of the fits have approximately 90% contribution from the faster time constant in the major groove, indicating that the dynamical

relaxation of water in the major groove follows mostly single exponential decay. Also, size of the site (in terms of no. of base pairs) and position or constituent of the site do not influence the MRT significantly as the overall MRT lies mostly between 61 to 65 ps for DD major groove sites.

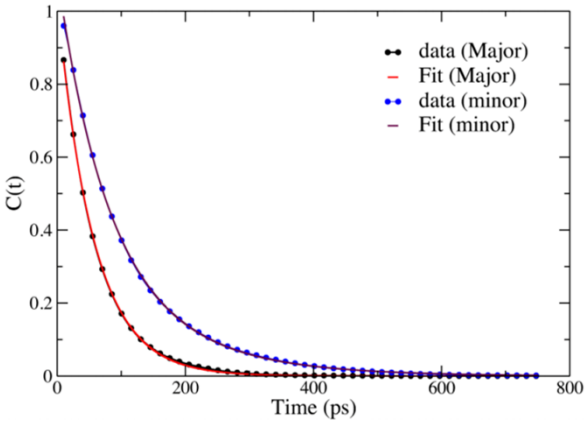

Figure 3.3 Correlation function and exponential fit for one representative calculation in a major and minor groove site. First, 10 ps data has been removed from the correlation function before fitting as done for the rest of the correlation functions.

On the other hand, water in the minor groove shows biexponential relaxation with a significant contribution for the larger time scale. Here, overall MRT is similar when four or more base pairs are considered as a site. For the sites involving two base pairs, the values are slightly smaller. The largest MRT value 128.2 ps is obtained for the minor groove site where four base pairs (A5 to T8) are considered as the site. X-ray crystallography studies show that A5 to A8 portion of DD has a narrow minor groove with a strong "spine of hydration".[57] The occupancy of water molecules in the grooves of DNA resembles that of water in the crystal structure. Therefore, the largest MRT for this region may be a consequence of that. However, we will see later that there are other effects also involved.

Table 3.1: Variation in the residence time with respect to the number of DNA base pairs in major and minor Grooves.

Major groove					Minor groove		
No. of base pairs	Base pair position	τ_1 $(Amplitude)$	τ_2 $(Amplitude)$	Total RT (ps)	τ_1 $(Amplitude)$	τ_2 $(Amplitude)$	Total RT (ps)

Two	2 to 3	49.9 (94.6%)	191.0 (5.4%)	57.5	62.9 (81.7%)	153.9 (18.3%)	79.6
	6 to 7	61.1 (91.1%)	124.3 (8.9%)	66.7	65.9 (78.1%)	184.8 (21.9%)	92.0
	10 to 11	50.0 (90.5%)	154.7 (9.5%)	59.9	53.4 (83.5%)	154.4 (16.5%)	70.1
Four	2 to 5	51.6 (94.4%)	176.7 (5.6%)	58.6	64.7 (74.9%)	197.7 (25.1%)	98.1
	5 to 8	57.7 (92.9%)	147.5 (7.1%)	64.0	86.6 (66.2%)	209.7 (33.8%)	128.2
	8 to 11	51.7 (89.5%)	143.7 (10.5%)	61.4	83.6 (65.6%)	160.6 (34.4%)	96.9
Six	2 to 7	52.4 (91.1%)	143.1 (8.9%)	60.4	68.1 (72.6%)	194.5 (27.4%)	102.7
	4 to 9	56.1 (93.1%)	149.6 (6.9%)	62.6	79.0 (70.3%)	200.7 (29.7%)	115.3
	6 to 11	53.2 (89.7%)	138.2 (10.3%)	61.9	68.6 (68.8%)	178.9 (31.2%)	103.0
Eight	2 to 9	53.3 (92.7%)	156.3 (7.3%)	60.8	69.6 (72.9%)	190.6 (27.1%)	102.4
	3 to 10	56.3 (94.0%)	177.4 (6.0%)	63.6	75.4 (74.8%)	195.9 (25.2%)	105.8
	4 to 11	52.8 (90.6%)	139.9 (9.4%)	61.0	69.4 (70.9%)	186.4 (29.1%)	103.5

Our calculated MRT for the minor groove of DD for A5T8 base pair is 128.2 ps, whereas the reported NMRD result of Denisov et al.[18] at similar temperature 27 °C is 200 ps at 27 °C. Although our value is somewhat smaller, the time scale obtained here is in the same order. Similarly, for the major groove of the same system, our result of 64 ps is smaller than the experiment (140 ps for DD at 27 °C).[18] Most of the other NMR studies including this one were performed at lower temperature, providing much longer residence time (>500 ps).[15, 18, 21] NMR studies have a lesser resolution at higher temperatures. Extrapolation of the results by Phan et al.[21], calculated at various lower temperatures in the minor groove of DD, to 298 K gives 100 ps, in good agreement with our present estimate. Although we find similar result for different base pairs, a clear signature of slower water dynamics in both major and minor grooves is present where the site is chosen to be middle for a set of base pairs (i.e., A6T7 for two base pair site, A5-T8 for four base pair site, and so forth). This phenomenon is consistent with NMR observation.[18, 58] Either sequential[21] or random placement[58] of AT base pairs shows that water near the middle of the DNA has longer time scale. Table 1 also shows that the values for different sizes of the reactant region are similar. Therefore, the effect of the base pair may be local. Water may escape quickly to the bulk before it samples the whole DNA. However, calculations of MRT on one system is not sufficient to decouple these different types of effects (i.e., chemistry, topography, and position in a DNA) on the water dynamics.

3.3.2 Distribution of Mean Residence Time

To decipher which effects are important and also to understand the range of variability of MRT of water around DNA, we calculated MRT of water, standardized through DD system as

mentioned above, onto the rest of the 20 different DNA systems around dinucleotide steps (2 base pairs). Because all DNA structures are taken to be 12 base pairs long, there are 9 sites in each groove leaving the terminal base pairs. Thus, there are altogether 360 sites providing us with 360 correlation functions.

Interestingly, the fitting did not always result in two time scales. Quality of fitting was decided on the basis of both χ^2 and correlation coefficient. We ensured that all the fit results had a low χ^2 and high correlation coefficient. The median of all 360 χ^2 is 0.04, and the correlation coefficient is 0.99996, indicating extremely good fit for all the correlation functions.

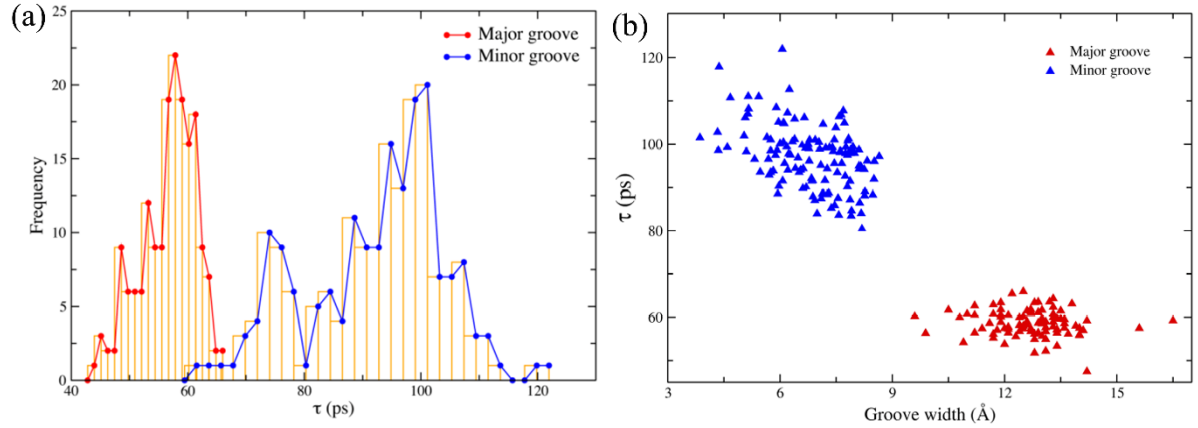

Figure 3.4 (a) Distribution of residence time, which shows a bimodal character based on the sites in major and minor grooves. (b) MRT values plotted against groove width of base pair steps (groove widths of some base pairs could not be calculated using Curves+62). Correlation between the groove width and MRT is clear as the values for a particular groove cluster together. Minor groove being narrower results in higher MRT of water and vice versa for the major groove.

Figure 3.4(a) shows the distribution of the overall MRT for all the 360 sites. The distribution is bimodal where MRTs of the major and minor groove sites fall under separate ones with only a little overlap. The difference in the time scale observed for major and minor groove can be attributed to the topography of the DNA. To illustrate this, the MRT values have been plotted with respect to the groove width of the DNAs calculated using Curves+[59] and are shown in Figure 3.5(b). From the figure, it can be seen that the time scales from different grooves fall in two different regions with non-overlapping groove widths. The minor groove has longer MRT compared to major groove in agreement with the experimental observation of a more rigid minor groove hydration by Jacobson et al.18 The values indicate that the A5-T8 region of DD shown before lies near the upper bound of the time scales of water near DNA (at least among the structures studied here). It is likely that the rigid DNA does not allow a drastic variation in the time scale found for the case of proteins.[29]

While it is believed that the differences in groove width and also electrostatic interaction in major and minor groove is the origin of grossly different time scales between the two grooves, we were interested to understand the origin of heterogeneity of time scales even in the same groove (either major or minor). As we have already seen that A5T8 of DD produces longest time scale, we can assume that either groove width variation or the difference in electrostatics of the base pairs (AT vs GC) or both may affect the time scale. However, as we will see, there is more to the story.

3.3.3 Origin of Heterogeneity

To investigate the origin of heterogeneity in the time scale of water in the same groove, we plotted MRT against different dinucleotide steps and also the position of those dinucleotide steps in Fig. 3.5. Figure 3.5a,b show the overall MRT against different dinucleotide steps. It is apparent from the figure that there is no correlation among the major groove sites. However, for the minor groove sites, AA, AT, TA base pairs have relatively higher MRT (slower relaxation) than GG, GC. It is known that AT base pairs have a stronger electrostatic effect compared to GC base pairs.[60-61] The electrostatics will have a stronger effect in the minor groove due to the narrow width of the groove. Therefore, we may attribute the slower relaxation to the chemistry of the base pair. Next, we plotted MRT values against base pair position alone in Figure 3.5c,d for major and minor groove sites, respectively.

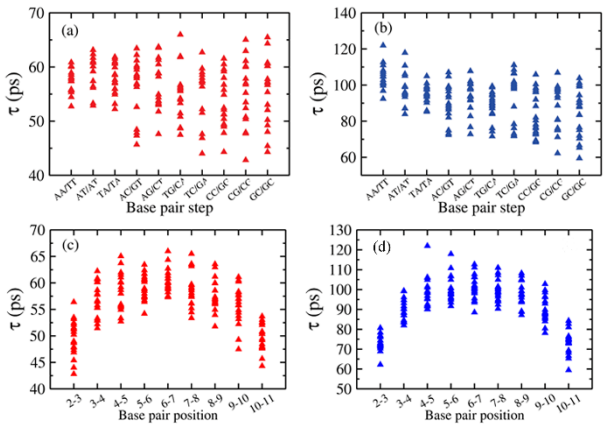

Figure 3.5 Variation of MRT for major (left panels, a and c) and minor (right panels, b and d) grooves plotted against base pair steps (top row) and base pair position along the DNA (bottom row).

From the figure, a strong and universal correlation across different chemical identity or topographic signatures becomes apparent based only on the position of the base pair in a DNA. As the position of the base pair goes toward the terminal, MRT decreases quadratically. This result is consistent with experiments, which showed that water around middle base pair has

longer relaxation time.36 NMR experiments by Leporc et al. using d(CTACTGCTTTAG)$_2$ also observed that AT in the middle has longer time scale.[58] We have also seen the same for our standardizing system, DD. However, DD contains AT group in the middle, and it has a narrow minor groove as well. Therefore, it was not possible to use only one DNA system to decipher this universal behavior. Because all three effects combined into the minor groove of A5-T8 in DD, the time scale obtained was highest among all the systems studied here (128.2 ps). The effect of the chemical nature of the base pair and position of the base pair in a DNA is, however, difficult to make quantitatively.

There is one more factor that may influence the water dynamics around DNA grooves (i.e., topography). It is probably the major reason for the bimodality observed in the distribution (Figure 3.4a). A water molecule in a deep and narrow minor groove will face difficulty to escape to bulk compared to other water molecules surrounding it. Wider and shallower major groove will pose less difficulty for the water to escape. To understand how DNA structure influences water dynamical properties in a particular groove, we have calculated various DNA parameters using Curves+, a method proposed by Lavery and co-workers.[59] The values (not shown here) indicated that major structural driving factors are groove width and depth for DNA. However the structural features observed for DNA do not show significant correlation with the observed MRT values, indicating that structural features of DNA are not the key factor in governing the MRT values. Topography will play a role only if the topographic difference (as in major and minor grooves) is significant.

3.3.4 Kinetic Model for MRT around DNA

To understand why position of the site in a DNA is the major driving factor for the heterogeneity within a particular (major or minor) groove, we propose a kinetic model, as shown in Figure 3.6. This model proposes the escape of water from DNA to bulk through different channels. From a site in the DNA in its first solvation shell, A_i, water can escape to the second solvation shell, B_i, with a rate constant k_{i12}. The reverse process of returning from the second to the first solvation shell is also allowed with rate constants $k_{-i12.}$ The escape from the second solvation shell to bulk is considered an irreversible process with a rate constant k_i. We also allow the movement of water along the DNA in the first and second solvation shells with rate constant $k_{i,i+1}$ and $k_{2i,i+1}$, respectively. Since there is no directional preference for going up or down the DNA, rate constants are taken to be equal. From the terminal base pair, A_t, an additional channel of escaping directly to bulk, denoted as P, is allowed. Therefore, this model is an extension of typical bound, quasi-bound, and free-states chosen for water escape. The only additional complexity comes from the fact that water can move along the DNA.

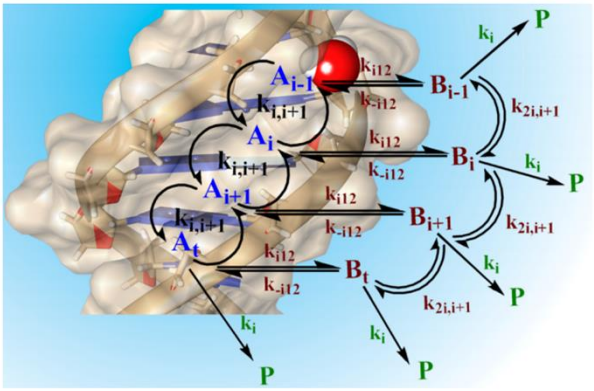

Figure 3.6 Schematic kinetic profile for water escape to bulk from DNA. A_i's and A_t's are non-terminal i-th base pair and terminal sites in the DNA, respectively. B_i's correspond to site near i-th base pair in the second solvation shell, and P denotes bulk. $k_{i,i+1}$ and $k_{2i,i+1}$denote rate constants for moving along the DNA in the first and second solvation shell corresponding to movement from i-th to i+1-th base pair, respectively. k_{i12} and k_{-i12} are the rate constants for water moving from first to second solvation shell and backward process for the i-th base pair, respectively. k_i is the rate constant for an escape to bulk from the second solvation shell. For simplicity, we used the same value for every base pair, thus five different rate constants.

Most of the processes mentioned above have a little barrier. Water movement from the first to second solvation shell shows highest barrier of ~1.4 kcal/mol (see Fig. 3.1d). Therefore, all kinetic steps are considered to be diffusion-limited, and the rate constants are calculated based on the expression of mean first passage time (with zero barrier) using the following relation,[62]

$$\tau(x) = \frac{1}{D}\int_{x}^{x_{max}} dy \int_{\text{xmin}}^{y} dz, \qquad (3.6)$$

Where $\tau(x)$ is the mean first passage time (MFPT) for water at the position x. There is a reflecting boundary at x_{min} and an absorbing boundary at x_{max}. Rate constants $k(x)$ is just the inverse of MFPT, $k(x) = \tau(x)^{-1}$. Calculations of the rate constant thus require diffusion coefficients and position of the two boundaries.

We calculated diffusion coefficients of water in the first and second solvation shell along the DNA and perpendicular to it from the local and directional VACF (see Section 3.2.4 for detail). For comparison, we also calculated water diffusion coefficient in the third solvation shell and bulk, as well. Figure 3.7 shows the diffusion coefficients parallel and perpendicular to the DNA in different solvation shells in the major and minor grooves. The calculations are performed using major and minor groove atoms of eight base pairs (C3 to G10) of DD. Note that, up to the third solvation shell, the diffusion coefficient is much smaller compared to bulk. The bulk diffusion coefficient matches with reported values obtained from simulation of TIP3P model[63]. The difference in the diffusion coefficient between major and minor grooves is also small.

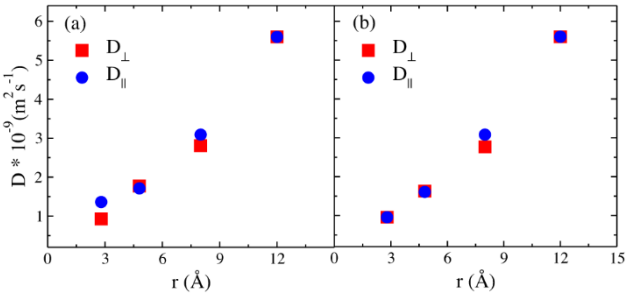

Figure 3.7 Diffusion coefficient parallel and perpendicular to DNA at different solvation shells including bulk for the (a) major and (b) minor groove. d_{min} is the minimum distance of a water molecule from DNA grooves. For the major groove, the first, second and third solvation shells are taken to be d_{min} < 3.1 Å, 3.1 Å < d_{min} < 6 Å, and 6 Å < d_{min} < 10 Å, respectively. For the minor groove, the first, second, and third solvation shells are defined as d_{min} < 3.4 Å, 3.4 Å < d_{min} < 6.6 Å, and 6.6 Å < d_{min} < 10 Å, respectively. The bulk value was calculated from a separate simulation with only water. Note that the difference between first and second solvation shell diffusion coefficient is small, indicating the effect of DNA on both the solvation shells.

Table 2 provides the diffusion coefficients, boundary values, and the rate constants calculated using the using Eq. 3.6. Note that, while x is the position of the minimum, reflecting boundary x_{min} is placed at the position of the preceding barrier whereas absorbing boundary is at the position of the next minimum. Base pair separation is considered to be 3.4 Å. Therefore, taking the relative position of x at 0.0, we have put the reflecting boundary at -1.7 Å (half of a base pair step), while the absorbing boundary is placed at the center of the next base pair (3.4 Å). The separation between boundaries is larger along the DNA compared to the solvation shell (Table 3.2). This results in a smaller rate constant for going along the DNA, $k_{i,i+1}$, compared to that for the escape from the first to the second solvation shell, k_{i12}, even though the diffusion coefficient is smaller along the perpendicular direction. Rate constant along the DNA in the second solvation shell, $k_{2i,i+1}$, has the same boundary except that the diffusion coefficient here is slightly higher, giving rise to a higher rate constant. For the same reason, the rate constant for second to the first solvation shell (backward reaction), k_{-i12}, is higher than the forward one. Finally, the rate constant for escape from the second solvation shell to bulk, k_i, is high due to again larger boundary separation. Note that the primary difference between minor and major grooves are in the rate constants for water movement between first and second solvation shell. This happens due broader solvation shell, which makes the boundary between two solvation shells less sharp.

Table 3.2: Local and directional diffusion coefficients, positions of the boundaries, and the rate constants used in the kinetic model.

Major groove				
D ($\times 10^{-9}$) ($m^2 s^{-1}$)	x_{min} (Å)	x(Å)	x_{max}(Å)	Rate Constants ($\times 10^{12}$) (s^{-1})

1.36	-1.7	0.0	3.4	$k_{i,i+1}$	0.012
1.72	-1.7	0.0	3.4	$k_{2i,i+1}$	0.015
0.93	2.0	2.8	3.6	k_{i12}	0.099
1.77	4.7	3.6	2.8	k_{-i12}	0.147
1.77	3.2	3.6	6.0	k_i	0.046
Minor groove					
D ($\times 10^{-9}$) (m^2s^{-1})	x_{max}(Å)	x(Å)	x_{min}(Å)	Rate Constants ($\times 10^{12}$) (s^{-1})	
1.10	-1.7	0.0	3.4	$k_{i,i+1}$	0.010
1.61	-1.7	0.0	3.4	$k_{2,i,i+1}$	0.014
0.96	2.0	2.8	4.8	k_{i12}	0.027
1.63	6.0	4.9	2.8	k_{-i12}	0.036
1.63	3.5	4.9	6.6	k_i	0.043

Using this simple kinetic model and five general rate constants, we calculated the water residence time, rather the time required for a water molecule to go to bulk (formation of P) starting at a particular i-th base pair. Figure 3.8 shows the timescale obtained from the fit to the time dependence of formation of P against the position of the base pair. This figure shows the similar trend of quadratic decrease from the middle position of the DNA towards the terminal as obtained from simulation with SSP approach (Fig. 3.5c, 5d). Interestingly, not only the qualitative behavior, we also find remarkable quantitative agreement between the values obtained for the major and minor grooves using the kinetic model and simulations (Fig. 3.5c and 3.5d). The highest value obtained from kinetic model for the middle of the minor groove is 93.3 ps while it is 92.0 ps for middle two base pairs of DD (Table 3.1) from simulation. For the major groove, the middle of the DNA provides 64.1 ps while we obtained 66.7 ps from the simulation for the middle of DNA (Table 3.1). The agreement of our kinetic model with simulation and experiment indicates that water diffusion pathway is not only perpendicular to DNA as assumed earlier[27]. The position dependent effect comes primarily from the diffusion of water along the DNA.

Therefore, only five general rate constants and simple diffusion limited kinetic model can explain the heterogeneous behavior of water dynamics around DNA. It is, however, possible to calculate the rate constants for a particular base pair, which will give us the detailed effect of chemistry and topography. Therefore, given a particular groove structure, the primary reason for heterogeneity is the position of the DNA rather than either chemistry or topography, which do play a role, albeit in smaller magnitude. For example, an exact estimation of rate constant in the narrow region of DD (A5-T8) gives the timescale as 140.8 ps for the minor groove and 76.9 ps for the major groove.

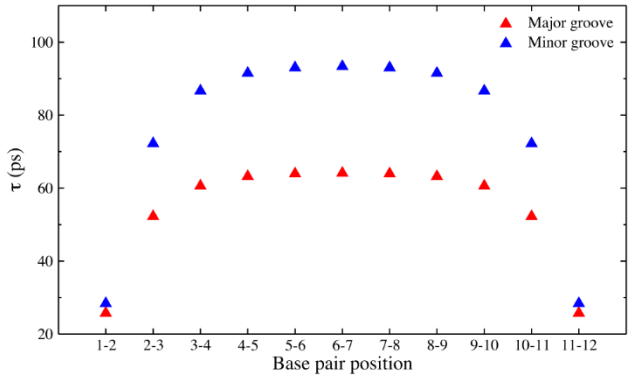

Figure 3.8 Timescale obtained from the kinetic model for sites in the major and minor groove along base pair position.

3.4 Conclusion

In this work, we addressed the origin of heterogeneity in the residence time of water around DNA base pairs using simulation and a kinetic model. In the simulation, the mean residence time was calculated using stable state picture approach, robustness and flexibility of which helped capturing the timescale in reasonable agreement with experiments. We calculated MRT of water around all ten different dinucleotide steps, each of which were placed randomly at different positions and different groove structures. This helped us to investigate the role of topography (groove widths and depths) and chemistry (base pair step). We showed that the distribution of MRT is bimodal where minor and major groove water molecules fall under separate distribution with little overlap. This indicates that groove structures do play a role. However, given a particular groove structure, the most dominating factor is the location of base pairs in a DNA near which the water molecule resides. Water in the middle of the DNA shows longest relaxation time that decreases quadratically towards the terminals. A simple kinetic model, where water is allowed to move along the DNA, captures not only the qualitative trend observed in the simulation and experiment but also agrees quantitatively. This indicates that the rate constants, calculated through local direction diffusion coefficients, are reasonable. Also, our model suggests that even in a homopolymer with same groove structure, water dynamics will be heterogeneous. Our kinetic model also shows that the difference in water timescales in major and minor groove is due to the solvation structure that affects the diffusion between the first two solvation shells. We believe that this work gives the water dynamic study around DNA a new perspective and throws some light on the governing factors and the origin of heterogeneity.

While this part of the study provides evidence that chemistry of the base pairs has lesser role in governing the dynamics of its surrounding solvent molecules, the question arises what happens when the effect of base pair chemistry is looked at by keeping different DNA base pairs in the middle of the DNA sequences. This has been investigated in the next chapter of the thesis.

3.5 References

1. Wang, L.; Yu, X.; Hu, P.; Broyde, S.; Zhang, Y. *J. Am. Chem. Soc.* **2007,** *129* (15), 4731-4737.
2. Yamasaki, K.; Akiba, T.; Yamasaki, T.; Harata, K. *Nucleic Acids Res.* **2007,** *35* (15), 5073-5084.
3. Jayaram, B.; Jain, T. *Annu. Rev. Biophys. Biomol. Struct.* **2004,** *33* (1), 343-361.
4. Russo, D.; Murarka, R. K.; Copley, J. R. D.; Head-Gordon, T. *J. Phys. Chem. B* **2005,** *109* (26), 12966-12975.
5. Modig, K.; Liepinsh, E.; Otting, G.; Halle, B. *J. Am. Chem. Soc.* **2004,** *126* (1), 102-114.
6. Otting, G.; Liepinsh, E.; Wuthrich, K. *Science* **1991,** *254* (5034), 974.
7. Denisov, V. P.; Halle, B. *Faraday Discuss.* **1996,** *103* (0), 227-244.
8. Grebenkov, D. S.; Goddard, Y. A.; Diakova, G.; Korb, J.-P.; Bryant, R. G. *J. Phys. Chem. B* **2009,** *113* (40), 13347-13356.
9. Pal, S.; Maiti, P. K.; Bagchi, B. *J. Chem. Phys.* **2006,** *125* (23), 234903.
10. Sen, S.; Andreatta, D.; Ponomarev, S. Y.; Beveridge, D. L.; Berg, M. A. *J. Am. Chem. Soc.* **2009,** *131* (5), 1724-1735.
11. Furse, K. E.; Corcelli, S. A. *J. Am. Chem. Soc.* **2008,** *130* (39), 13103-13109.
12. Zhong, D.; Pal, S. K.; Zewail, A. H. *Chem. Phys. Lett.* **2011,** *503* (1), 1-11.
13. Furse, K. E.; Corcelli, S. A. *J. Am. Chem. Soc.* **2011,** *133* (4), 720-723.
14. Pal, N.; Verma, S. D.; Sen, S. *J. Am. Chem. Soc.* **2010,** *132* (27), 9277-9279.
15. Liepinsh, E.; Otting, G.; Wüthrich, K. *Nucleic Acids Res.* **1992,** *20* (24), 6549-6553.
16. Jacobson, A.; Leupin, W.; Liepinsh, E.; Otting, F. *Nucleic Acids Res.* **1996,** *24* (15), 2911-2918.
17. Sunnerhagen, M.; Denisov, V. P.; Venu, K.; Bonvin, A. M. J. J.; Carey, J.; Halle, B.; Otting, G. *J. Mol. Biol.* **1998,** *282* (4), 847-858.
18. Denisov, V. P.; Carlström, G.; Venu, K.; Halle, B. *J. Mol. Biol.* **1997,** *268* (1), 118-136.

19. Jóhannesson, H.; Halle, B. *J. Am. Chem. Soc.* **1998,** *120* (28), 6859-6870.

20. Halle, B.; Denisov, V. P. *Biopolymers* **2000,** *48* (4), 210-233.

21. Phan, A. T.; Leroy, J.-L.; Guéron, M. *J. Mol. Biol.* **1999,** *286* (2), 505-519.

22. Luise, A.; Falconi, M.; Desideri, A. *Proteins Struct. Funct. Bioinf.* **2000,** *39* (1), 56-67.

23. Makarov, V. A.; Andrews, B. K.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **2000,** *79* (6), 2966-2974.

24. Brunne, R. M.; Liepinsh, E.; Otting, G.; Wüthrich, K.; van Gunsteren, W. F. *J. Mol. Biol.* **1993,** *231* (4), 1040-1048.

25. Rocchi, C.; Rita Bizzarri, A.; Cannistraro, S. *Chem. Phys.* **1997,** *214* (2), 261-276.

26. García, A. E.; Stiller, L. *J. Comput. Chem.* **1993,** *14* (11), 1396-1406.

27. Bonvin, A. M. J. J.; Sunnerhagen, M.; Otting, G.; van Gunsteren, W. F. *J. Mol. Biol.* **1998,** *282* (4), 859-873.

28. Roy, S.; Bagchi, B. *J. Phys. Chem. B* **2012,** *116* (9), 2958-2968.

29. Sterpone, F.; Stirnemann, G.; Laage, D. *J. Am. Chem. Soc.* **2012,** *134* (9), 4116-4119.

30. Henchman, R. H.; McCammon, J. A. *Protein Sci.* **2009,** *11* (9), 2080-2090.

31. Hua, L.; Huang, X.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2006,** *110* (8), 3704-3711.

32. Schröder, C.; Rudas, T.; Boresch, S.; Steinhauser, O. *J. Chem. Phys.* **2006,** *124* (23), 234907.

33. Beck, D. A. C.; Alonso, D. O. V.; Daggett, V. *Biophys. Chem.* **2002,** *100* (1), 221-237.

34. Yonetani, Y.; Kono, H. *Biophys. J.* **2009,** *97* (4), 1138-1147.

35. Yonetani, Y.; Kono, H. *Biophys. Chem.* **2012,** *160* (1), 54-61.

36. Laage, D.; Hynes, J. T. *J. Phys. Chem. B* **2008,** *112* (26), 7697-7701.

37. Northrup, S. H.; Hynes, J. T. *J. Chem. Phys.* **1980,** *73* (6), 2700-2714.

38. Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981,** *78* (4), 2179-2183.

39. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005,** *26* (16), 1668-1688.

40. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

41. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007,** *92* (11), 3817-3829.

42. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008,** *4* (3), 435-447.

43. Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005,** *88* (4), 2472-2493.

44. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

45. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

46. Nosé, S. *Mol. Phys.* **1984,** *52* (2), 255-268.

47. Hoover, W. G. *Phys. Rev. A* **1985,** *31* (3), 1695-1697.

48. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

49. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

50. Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980,** *102* (13), 4287-4294.

51. Lounnas, V.; Pettitt, B. M. *Proteins Struct. Funct. Bioinf.* **1994,** *18* (2), 133-147.

52. Makarov, V. A.; Andrews, B. K.; Pettitt, B. M. *Biopolymers* **1998,** *45* (7), 469-478.

53. Chandler, D. *Introduction to Modern Statistical Mechanics*. Oxford University Press: New York, 1987; Vol. 1.

54. Hansen, I. R. M. J. P. *Theory of Simple Liquids*. Academic Press: San Diego, 1990.

55. Kulkarni, M.; Mukherjee, A. *J. Chem. Phys.* **2013,** *139* (15), 155102.

56. Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981,** *151* (3), 535-556.

57. Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983,** *163* (1), 129-146.

58. Leporc, S.; Mauffret, O.; Antri, S. E.; Convert, O.; Lescot, E.; Tevanian, G.; Fermandjian, S. *J. Biomol. Struct. Dyn.* **1998,** *16* (3), 639-649.

59. Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009,** *37* (17), 5917-5929.

60. Jayaram, B.; Sharp, K. A.; Honig, B. *Biopolymers* **1989,** *28* (5), 975-993.

61. Lavery, R.; Pullman, B. *Nucleic Acids Res.* **1981,** *9* (24), 7041-7051.

62. Zwanzig, R. *Nonequilibrium Statistical Mechanics*. Oxford University Press: New York, 2001.

63. Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000,** *114* (1), 363-366.

Chapter 4

Water Mediated Ultraslow Dynamics of Hydrated Ionic Liquids near CG Rich DNA: Consequence to DNA Stability

4.1 Introduction

Once a sole genetic information carrier agent, DNA has recently found its new uses in developing advanced materials[1] and in molecular computing devices[2], as molecular motors[3] and in hybrid catalysis[4-5] with promising utility as biosensors and in asymmetric synthesis. Although the structure of DNA has been considered to be stable, the long-term stability of DNA has attracted attention with the advancement of different materials. DNA structure is susceptible to denaturation with the variation in temperature, pH, and ionic strength. At physiological pH, DNA is found to be stable in water only for a few days for its tendency to undergo hydrolytic reactions.[6] As an alternate medium, ionic liquids (ILs) are increasingly used as a storage material for DNA and other biomolecules because of their low vapour pressure, non-toxicity, and high chemical and thermal stability.[7-8] The observed long term storage ability of DNA in hydrated ILs[9] have led to studies that unveil the interactions responsible for this stabilization.

The CG base pairs are known to be more stable compared to AT base pairs in aqueous media not only owing to its three Hydrogen bonds compared to two in AT, but also due to the secondary hydrogen bonding (HB)[10] and stacking interactions.[11] A reversal of this trend has been reported by von Hippel *et al.* with alkyl ammonium ions, where the stability was found to reduce with increasing the CG content.[12-13] The hydrated choline dihydrogen phosphate solution has also been shown to impart similar stability to AT base pairs compared to sodium chloride solutions through melting temperature study.[14]Several studies have used experimental and computational methods or a combination of both to unravel interactions between DNA base pairs and the cations of ILs[15-19], where AT base pairs are found to be more stable than CG in the presence of IL. The enhanced stabilization of AT base pairs has been attributed to the preferential binding of IL cation with more electrostatic AT bases[18], whereas the general stabilization for DNA in ILs has been ascribed to the electrostatic interactions between the cations and the phosphate groups of DNA[19]. The favourable entropic change due to the reported change in the hydration pattern in presence of ILs[20] also has been considered to be responsible for the enhanced stability of DNA.[21]

The structure and dynamics of ILs in water has been the focus of several studies. Although the mechanisms and interactions responsible for DNA stabilization in ILs have been studied extensively, *the dynamics of these ions around DNA has been largely ignored*. The ILs are known to be viscous, and the diffusion coefficients are few orders of magnitude lower than water even at elevated temperatures.[22] Therefore, the dynamics of the constituent cations is expected to be even slower near the negatively charged DNA.

Here, we have focused on the dynamics of ILs through the mean residence time (MRT) of different IL cations near various DNA sites using all-atom molecular dynamics simulation. Simultaneously the MRT for water was monitored to probe the dynamical relaxation of biological water.[23] We have used a recent and rigorous prescription of MRT calculation of Laage and Hynes[24] which measures the time taken for a cation to escape the solvation shell of DNA. This method has shown to produce MRT around different DNA base pairs in close agreement with NMR results.[25] Interestingly, we find that the dynamics of IL cations is not correlated with their strength of binding to the DNA. While IL cations form stronger interaction with AT than CG, the MRT is longer for the CG than AT. Calculations carried out with variation in anions showed little change in the observed pattern. DNA parameters also do not show the root of the cause. Our result will show that water, due to its stronger interaction with CG compared to AT, makes the IL dynamics also slower. We discuss how these dynamical behaviours can affect the binding propensity of IL cations with CG and AT base pairs, which is expected to have an additional unfavourable entropic contribution for CG base pairs.

4.2 Computation Details

4.2.1 System Design

We have used all-atom molecular dynamics simulations to probe the MRT of three different cations in five different IL combinations around two different DNA dodecamer sequences, 5'-CGCGCGCGCGCG-3' and 5'-CGCGAATTCGCG-3'. These dodecamers differ only in the middle four base pairs; one sequence has CGCG sequence (denoted as CG-DNA hereafter) while the other has AATT (referred to as AT-DNA hereafter) sequence. Since MRT depends on the position along the DNA[25], we have calculated MRT of IL and water around the major, minor groove and phosphate groups of the middle four base pairs to obtain only the effect of the DNA sequence on the dynamics of the IL.

We have chosen the following five different ILs: (i) 1-butyl-3-methylimidazolium chloride ([BMIM][Cl]), (ii) 1-butylpyridinum chloride ([BPYR][Cl]), (iii) choline chloride ([Choline][Cl]), (iv) 1-butylpyridinium hexafluorophosphate ([BPYR][PF_6]) and (v) 1-butylpyridinium tetrafluorobromide ([BPYR][BF_4]). The five ILs have three different cations, $[BMIM]^+$, $[BPYR]^+$, and $[Choline]^+$, containing structural variations to allow us to study the interaction of the base pairs with a five-membered ($[BMIM]^+$), six membered ($[BPYR]^+$), and alkylammonium (Choline) cation. The choice of the ILs is governed by the variation in the structure of the cations. Fig. 4.1 shows the chemical structure of the IL cations used in this study.

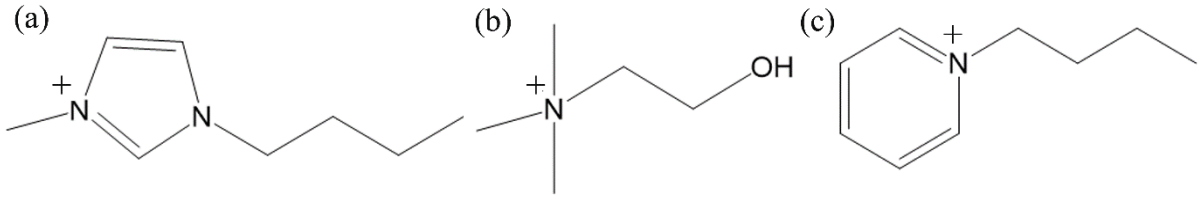

Figure 4.1 Chemical structures of (a) 1-butyl-3-methylimidazolium cation ($[BMIM]^+$), (b) choline cation and (c) 1-butylpyridinum cation ($[BPYR]^+$).

The solubility and stability of proteins have been previously shown in hydrated ionic liquids.[26]The composition of water and IL was taken from the study by Chandran *et al.*[16] for making 50 wt% mixture. Further simulations in 25 wt% of [BIMIM][Cl] and in pure water were also carried out for the two DNA systems where the number of IL and water were kept same for both the DNAs. A simulation for the sequence 5'-ATATATATATAT-3' was also done at 50 wt% [BIMIM][Cl] solution to compare the effect of the terminal residues in the dynamics of IL cations and water near AT base pairs.

4.2.2 Simulation Details

The DNA structures in the B-form were generated from the above sequences using the nucleic acid builder (NAB) program in AMBERTools[27], which was also used to generate the co-ordinate and topology files with AMBER99/parmbsc0 force field[28-29]. These coordinates and topologies were converted to GROMACS[30] format using amb2gmx.pl program[31]. For the ionic liquids, the parameters were taken from the study of Sambasivarao *et al.*[32]TIP3P water model[33] was used along with the ILs to solvate the dodecamers. An additional 22 IL cations were used to neutralize the system. All the systems were first energy minimised using steepest descent

method[34]. This was followed by heating up to 300K using Berendsen thermostat[35] with a coupling constant of 0.2 ps and a harmonic restraint of 250 kcal/mol/nm^2 on heavy atoms of the DNA. The restraints were reduced sequentially to 5000 cal/mol/nm^2 in seven simulations of 50 ps each at constant temperature of 300 K and constant pressure of 1 bar using Berendsen thermostat and barostat[35], respectively. Finally, a 110 ns long simulation was carried out for each system without any restraint at 300 K temperature and 1 bar pressure using Nose-Hoover thermostat[36-37] and Parrinello-Rahman barostat[38], respectively, with a coupling constant of 0.2 ps for each. PME electrostatics[39]with a 1 nm cut-off was used for the electrostatic interaction. Same cut-off distance was used for the van der Waals interactions, too. Leaving the first 10 ns simulation as equilibration, analyses were performed on the last 100 ns of the simulation to calculate MRT and other properties. All the simulations were carried out using GROMACS[30] molecular dynamics software. For the simulation in pure water, we carried out 1 ns equilibration followed by a final production run for 10 ns in NPT conditions with the same parameters used for other simulations.

4.2.3 Mean Residence Time (MRT) Calculation

The method for MRT calculation is discussed in Section 2.2 of Chapter 2 of the thesis. The other details on the MRT calculations have been discussed in Section 3.2.3 of Chapter 3. Since DNA is not a spherical molecule, the RDF gives only a rough picture of the solvation shells. Therefore, in the previous chapter, proximal radial distribution function (PRDF) has been used to obtain information on solvation shells. However, here we intend to compare the MRT around AT and CG base pairs for a particular IL-water solution. We do not aim to compare the MRT for different ILs in this study. Since PRDF is computationally demanding, the conventional RDF may be sufficient in this case for the purpose of defining reactant and products. Here we have used the same distance cut off for a particular IL with both the DNA systems. Thus, although the absolute values of the MRT may depend on the specific distance cut-off, the relative magnitude will not depend significantly on the choice of the cut-off. The RDF and PMF plots obtained for the three different cations $[BMIM]^+$, $[Choline]^+$ and $[BPYR]^+$ near major groove atoms of the DNA systems are shown in Fig. 4.2. Reactant is defined to be the smallest distance below which the first prominent RDF peak is observed with a corresponding stable minimum observed in the PMF. Product is the distance above which there is no chance of recrossing, far away from the minima where PMF flattens out. The numerical values for the cut-off distances for all the different sites are given in Table 4.1. To define the product state, we consider the IL cation to be far away from the whole DNA instead of only the DNA site.

This way, the movement along the DNA still contributes to the overall MRT. Absorbing boundary conditions were used for the product state to prevent re-reaction (product to reactant conversion) as was prescribed[40] and used earlier[25, 40]. The vertical dashed line in Fig. 4.2 denotes the reactant and product for different ionic liquids around phosphate groups.

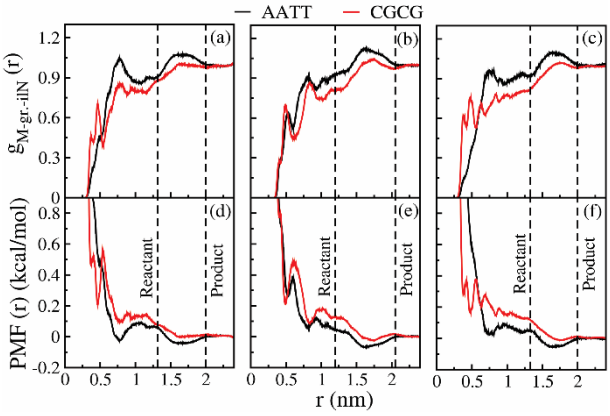

Figure 4.2 The RDF (top panel) and PMF (bottom panel) plots for major groove atoms of the DNAs with $[BMIM]^+$ (a and d), $[choline]^+$ (b and e) and $[BPYR]^+$ (c and f), respectively.

Table 4.1 The distance cut-off taken for calculation of MRTs of IL cations.

IL solution	Reactant state cut-off (nm)			Product state cut-off (nm)		
	Major	Minor	Phosphate	Major	Minor	Phosphate
[BMIM][Cl]	1.32	1.20	0.97	2	1.65	1.7
[Choline][Cl]	1.20	1.06	1.17	2.05	1.9	1.8
[BPYR][Cl]/ [BPYR][PF_6]/[BPYR][BF_4]	1.33	1.10	0.67	2	2	1.55

4.3 Results

4.3.1 Mean Residence Time of Cations around DNA.

The exponential function for the IL cations has a very similar pattern as shown in Fig. 3.3 in Chapter 3. Unlike the case of water, the correlation function here takes a much longer time to decay. Here also the correlation function can be fitted with a biexponential function as the following:

$$C(t) = C_1 e^{-t/\tau_1} + C_2 e^{-t/\tau_2}, \qquad (4.1)$$

Here C1 and C2 are the amplitudes of time constants τ_1 and τ_2, respectively. Similar to the case in Chapter 3, here also τ_1 is taken to be the faster, and τ_2 is taken to be the slower time scale by convention. The overall MRT values are obtained using the following relation:

$$\tau = \frac{C_1\tau_1 + C_2\tau_2}{C_1 + C_2}, \qquad (4.2)$$

The MRT values obtained for different IL cations are shown in Table 4.2. The values of C1, C2, τ_1 , τ_2 and the fitting parameter are given in Table 4.3.

Table 4.2 Mean residence time of IL cations around different sites of AT-DNA and CG-DNA.

	Cation MRTs (ns)					
	Major groove		**Minor groove**		**Phosphate**	
Ionic liquids	**AATT**	**CGCG**	**AATT**	**CGCG**	**AATT**	**CGCG**
[BIMIM][Cl]	6.8	77.0	8.2	50.9	7.2	25.3
[Choline][Cl]	17. 9	2.3×10^4	20.2	2.0×10^4	13.4	20.8
[BPYR][Cl]	4.9	79. 9	6.2	99.2	5.7	61.3
[BPYR][PF_6]	3. 5	8.1	4.1	11.1	3.6	8.1
[BPYR][BF_4]	3.7	7.1	4.4	9.7	4.4	9.2

Table 4.3 The coefficients, time scales and the fitting parameters for MRT calculation of IL cations.

IL solution	**DNA**	**Site**	**C_1 (%)**	**τ_1 (ns)**	**C_2 (%)**	**τ_2 (ns)**	**Correlation coefficient**	**χ^2**
[BMIM][Cl]	AATT	Maj	91.2	1.79	8.8	59.06	0.99	3.57
		Min	88.3	1.67	11.7	57.5	0.99	5.02
		Phs	90.2	1.37	9.8	60.35	0.99	3.25
	CGCG	Maj	85.2	10.16	14.8	463.23	0.99	7.99
		Min	79.3	9.8	20.7	207.82	0.99	2.24
		Phs	73.8	7.09	26.2	76.69	0.99	1.12
[Choline][Cl]	AATT	Maj	22.6	2.49	77.4	22.34	0.99	0.89
		Min	12.1	2.06	87.9	22.74	0.99	1.97
		Phs	40.9	1.17	59.1	21.92	0.99	0.47
	CGCG	Maj	85.1	15.32	14.9	1.56×10^5	0.99	27.19
		Min	81.8	16.14	18.2	1.1×10^5	0.99	34.52
		Phs	75	7.87	25	59.58	0.99	2.8
		Maj	89.9	3.16	10.1	20.22	0.99	1.37

[BPYR][Cl]	AATT	Min	90.3	3.89	9.7	27.43	0.99	6.84
		Phs	86	2.65	14	24.1	0.99	1.23
	CGCG	Maj	82	7.77	18	407.48	0.99	6.83
		Min	72.3	9.36	27.7	333.46	0.99	11.42
		Phs	79.9	5.76	20.1	282.83	0.99	2.99
[BPYR][PF_6]	AATT	Maj	82.6	2.28	17.4	9.1	0.99	0.97
		Min	83.8	2.83	16.2	10.76	0.99	3.77
		Phs	80.9	1.89	19.1	11.04	0.99	0.9
	CGCG	Maj	86.9	4.04	13.1	35.18	0.99	4.35
		Min	82.3	5.00	17.7	39.29	0.99	5.87
		Phs	76.6	2.58	23.4	26.28	0.99	3.06
[BPYR][BF_4]	AATT	Maj	88.6	2.21	11.4	15.33	0.99	0.54
		Min	86.6	2.52	13.4	16.77	0.99	4.41
		Phs	82.7	1.91	17.3	16.06	0.99	1.59
	CGCG	Maj	83.1	3.58	16.9	24.26	0.99	5.28
		Min	79.1	4.69	20.9	28.73	0.99	7.59
		Phs	77	2.47	23	31.93	0.99	6.78

Table 4.2 shows that MRT of IL cations are generally very high, in nanoseconds, owing probably to the strong electrostatic attraction between the negatively charged DNA and positively charged cations. Most of the correlation functions did not decay to zero. However, the bi-exponential fit of the correlation function provides a good measure of the MRT. The inclusion of polarizability in the force-field parameters for ILs can also influence the ILs' properties and the MRT values. Tsuzuki *et al.*[41] showed that the diffusion coefficients of IL solutions obtained from simulations can be an order of magnitude different compared to experimental values, although the experimental trend remained similar to the simulation results. Similarly in our calculations, the qualitative feature mentioned below is quite evident and is the focus of the present study. Table 4.2 indicates that the MRT values are always significantly higher for the CG-DNA compared to the AT-DNA for every IL and all the sites (major groove, minor groove, and phosphate group) studied here. Choline cations have the largest MRT out of the three cations considered in the study. MRT values become shorter for both the dodecamers when the anion is varied from a smaller Cl^- to larger PF_6^- or BF_4^- anions. The reduction is much higher for the CG-DNA compared to the AT-DNA. The MRT values were still found to be higher for CG-DNA compared to AT-DNA. The values were found to be similar with PF_6^- and BF_4^- indicating that the higher charge density of Cl^- ion may make the solution more viscous, resulting in slower dynamics of IL cations. Therefore, we will be

focusing only on the solutions where Cl^- is used as the anion, where we find a significant difference in the MRT values for AT and CG-DNA.

Our observation of a longer MRT near CG-DNA is supported by the NMR study of Portella *et al*.[17] where the MRT values have been determined using NMR technique for the cations within close contact of base protons. It was found that the major groove of DNA sequence GCGGGGCCCCGC has higher MRT values compared to sequences having higher AT content. Also, the MRT values have been reported to be longer in the minor groove compared to major groove. However, the timescale obtained from this experiment is much less compared to the present calculation due to the difference in the definition of the residence criteria. While experimental probe detect the MRT within 0.4 nm[17], here we probed for a complete removal of the IL cations or water from the DNA site.

Since our calculation probes into complete escape of the cation from the DNA sites, the values obtained here are much larger compared to the above where only the contact with DNA protons were probed. Interestingly, the qualitative trend is still preserved. To get an estimate of the error associated with the calculated MRT values, we have done error analysis using the AT-DNA in [BPYR][BF_4]solution as a representative one. We could not calculate the error in all the values because of the associated high computational cost. To estimate the error in the MRT values, the 100 ns simulation with [BPYR][BF_4] was divided into two 50ns segments and the C(t) was calculated for 25 ns. The reason for choosing [BPYR][BF_4] solution comes from the fact that in this case, the correlation function decays rapidly for the IL cation and hence in 25 ns, the C(t) can be expected to reach close to zero. Fitting the data with a bi-exponential function, the timescales have been calculated from which we have calculated the error as the standard deviation (S.D.) of MRT values using Eq. 4.3.

$$\text{S.D.} = \sqrt{((\tau_{100\text{ns}} - \tau^1_{50\text{ns}})^2 + (\tau_{100\text{ns}} - \tau^2_{50\text{ns}})^2)/2}, \qquad (4.3)$$

The errors associated with the MRT values are given in Table 4.4. The error associated with the MRT values of water present in the same solution are also given in Table 4.4. The values in the table show that the error associated with the MRT values are small for all cases.

Table 4.4 The timescales with different segment of trajectory and the error associated with the MRT values of $[BPYR]^+$ and water.

Solvent	Site	MRT (ns) from 100ns (τ_{100ns})	MRT (ns) from first 50ns (τ^1_{50ns})	MRT (ns) from last 50ns (τ^1_{50ns})	S. D. (ns)
	Maj	3.7	3.77	3.71	0.05

$[BPYR]^+$	Min	4.43	4.91	4.69	0.38
	Phs	4.36	4.65	4.49	0.22
Water	Maj	0.27	0.27	0.28	0.003
	Phs	0.18	0.18	0.19	0.008

4.3.2 Probing the Higher MRT Values near CG Base Pairs.

To understand the reason for higher MRT values near CG base pairs, the first and obvious thing to look for is the strength of interaction between IL cations and the DNA base pairs. Therefore, we have calculated both the van der Waals (vdW) and Coulombic interaction energies of the cations with the DNA sites consisting of either major groove atoms or minor groove atoms or phosphate atoms of the middle four base pairs. Fig. 4.3 shows the energies obtained for the two dodecamer systems studied here. From the figure, we see that AT interact strongly (lower energy) with DNA than CG in general. In some of the cases, both AT and CG show similar interaction. The stronger interactions in the minor grooves of AT base pairs have been reported previously[17-18].Among different sites of the DNA, interactions with IL cations are strongest with the negatively charged phosphate groups, as expected. Since the interaction is stronger with AT rather than CG, the heuristic explanation of the stronger binding for a higher MRT does not apply.

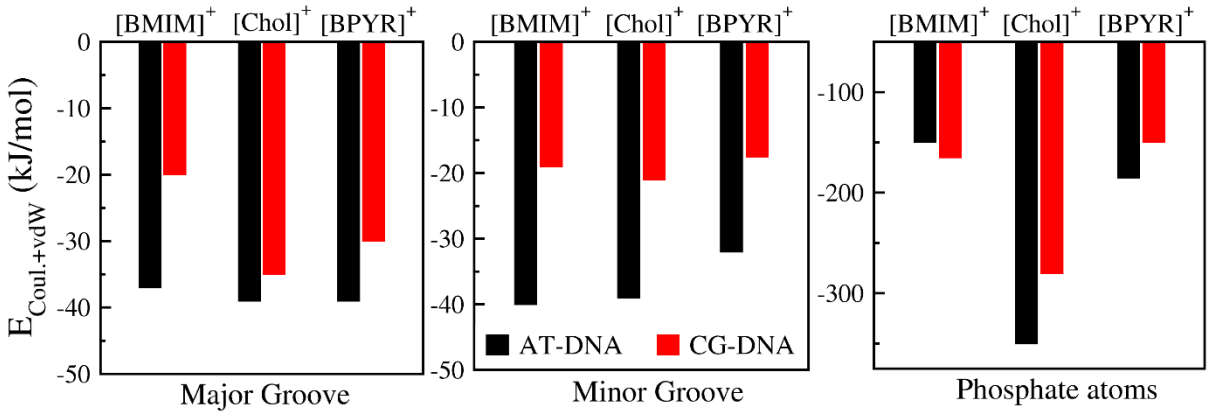

Figure 4.3 The interaction energies between $[BMIM]^+$, $[choline]^+$ and $[BPYR]^+$ with major groove (left panel), minor groove (middle panel) and phosphate group (right panel) atoms of the middle four base pairs of the two DNA systems considered. Red lines indicate the interaction with CG-DNA while black line indicates the interaction energies with AT-DNA.

To probe whether the DNA topography is responsible for the non-intuitive behaviour mentioned above, we have calculated DNA parameters for the two dodecamer systems in

different IL solutions using Curves+, a method proposed by Lavery and co-workers.[42] We looked at various DNA parameters such as groove width, groove depth, rise, roll, slide, shift, tilt and twist which might govern the dynamics of the ILs. The values (not shown here) indicated that although there are variations in the DNA parameters for the two dodecamers, there is no consistent alteration in the values that can give rise to slower dynamics near CG-DNA in all the ILs considered. Therefore, the dynamical pattern near CG-base pairs is unlikely to originate from the topographical variation of the two DNAs.

4.3.3 Water MRT near different sites

We then investigated the MRT of water for these IL-water solutions of dodecamers. First, we calculated the RDF of water and obtained the PMF to define the stable reactant and product states. The RDF and PMF profiles for water near the major groove in different IL solutions are shown in Fig. 4.4. A similar plot is obtained for water near phosphate atoms. The minor grooves of the dodecamers do not show distinguishable peaks in water RDF, may be because these sites are not accessible to water due to higher occupancy by IL cations. Hence, we only calculated the water MRT near major groove and the phosphate groups.

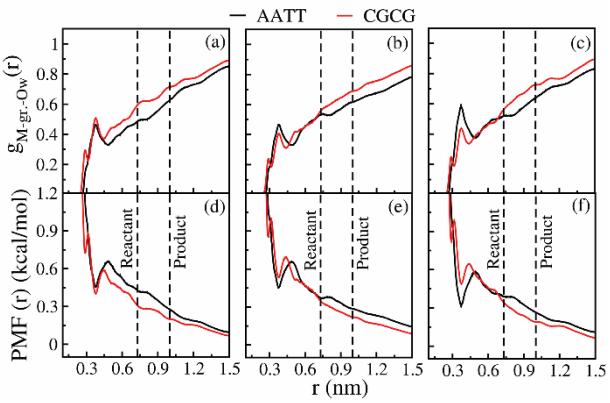

Figure 4.4 The RDF (top panel) and PMF (bottom panel) plots for major groove atoms of the DNAs with water oxygen in [BIMIM][Cl] solution (a and d), [Choline][Cl] solution (b and e) and [BPYR][Cl] solution (c and f), respectively.

We have tabulated the MRT values of water around major groove and phosphate groups of the middle four base pairs of the two dodecamers in Table 4.5. The different values of C1, C2, τ_1, τ_2 are given in Table 4.6. Note that, the values obtained here are higher compared to our previous study[25] which is discussed in Chapter 3 involving only water, where the largest value obtained was 0.064 ns for the major groove site. Here, the smallest value for the major groove

is itself 0.19 ns. This may be due to two reasons: (i) IL prevents the escape of water, or (ii) the RDF and hence the cut-off values are larger due to partial blockage of the sites by IL cations. Therefore, the water MRT may be affected both directly and indirectly by the ILs.

Water MRT also differs between AT and CG DNA in presence of IL cations. The MRT values of water are remarkably higher near CG-rich DNA compared to the AT-rich DNA in all the IL solutions. The energy calculations between water and DNA sites reveal now an opposite trend as was observed in case of the IL cations. Fig. 4.5 shows the energy from both vdW and Coulombic interaction between water and the major groove and phosphate group atoms. Here, water seems to interact much strongly with CG-rich DNA compared to AT-rich DNA. The interaction energy between water and the major groove atoms of the two dodecamer seem to be close, though.

Table 4.5 Water MRT values in different IL solutions.

	Water MRTs (ns)			
	Major groove		Phosphate	
IL solution	AATT	CGCG	AATT	CGCG
[BIMIM][Cl]	0.22	1.56	0.12	0.83
[Choline][Cl]	0.34	1.58	0.18	0.77
[BPYR][Cl]	0.23	1.14	0.14	0.64
[BPYR][PF_6]	0.19	0.46	0.11	0.29
[BPYR][BF_4]	0.27	0.59	0.18	0.35

Table 4.6 The MRT calculation results for water in different IL solutions. Sites containing major groove atoms and the phosphate atoms are denoted as Maj, Phs, respectively.

IL solution	DNA	Site	C_1 (%)	τ_1 (ns)	C_2 (%)	τ_2 (ns)	Correlation coefficient	χ^2
[BMIM][Cl]	AATT	Maj	89.3	0.15	10.7	0.84	0.99	0.06
		Phs	91.8	0.08	8.2	0.57	0.99	0.06
	CGCG	Maj	88.4	0.91	11.6	6.55	0.99	0.19
		Phs	23.3	2.15	76.7	0.42	0.99	0.31
[Choline][Cl]	AATT	Maj	82.0	0.21	18	0.83	0.99	0.08
		Phs	21.6	0.54	78.4	0.08	0.99	0.19
		Maj	95.1	0.91	4.9	14.74	0.99	0.23

	CGCG	Phs	75.9	0.42	24.14	1.89	0.99	0.28
[BPYR][Cl]	AATT	Maj	86.8	0.16	13.2	0.71	0.99	0.04
		Phs	84.5	0.08	15.5	0.47	0.99	0.09
	CGCG	Maj	88.5	0.77	11.5	3.96	0.99	0.25
		Phs	80.3	0.36	19.7	1.79	0.99	0.21
[BPYR][PF_6]	AATT	Maj	92.1	0.14	7.9	0.73	0.99	0.05
		Phs	84.1	0.06	15.9	0.35	0.99	0.07
	CGCG	Maj	100	0.46	0	0	0.99	1.84
		Phs	80.1	0.18	19.9	0.73	0.99	0.09
[BPYR][BF_4]	AATT	Maj	93.6	0.15	6.4	2.06	0.99	0.16
		Phs	92.4	0.08	7.6	1.4	0.99	0.37
	CGCG	Maj	92.6	0.42	7.4	2.69	0.99	0.21
		Phs	87.9	0.21	12.1	1.33	0.99	0.21

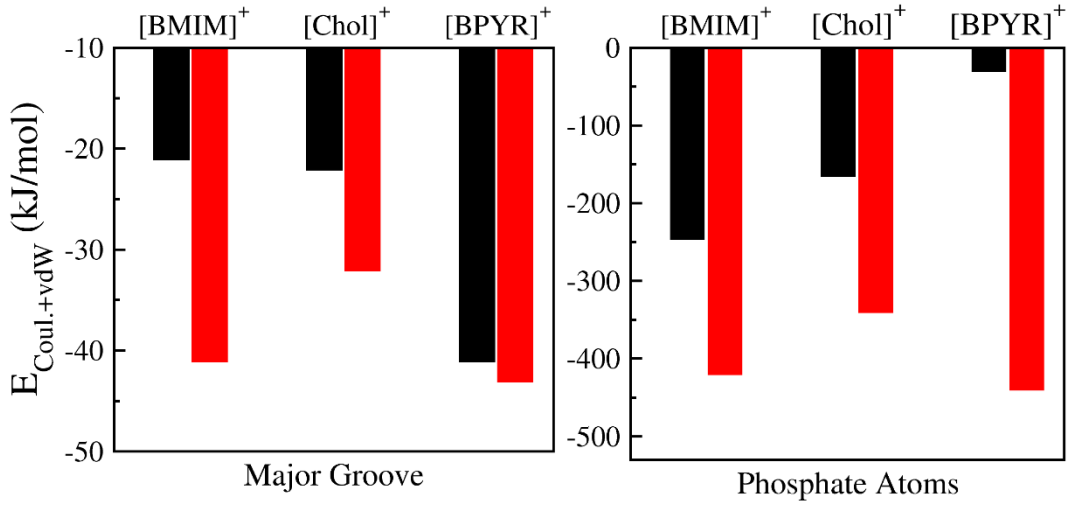

Figure 4.5 The interaction energies of water with major groove (left panel) and with phosphate group (right panel) atoms of the middle four base pairs in [BIMIM][Cl] solution, [Choline][Cl] solution and [BPYR][Cl] solution for the two DNA systems considered.

Therefore, although the longer MRTs of water near the CG base pairs could be attributed to the stronger interactions, the reason for longer MRT of IL cations is still not apparent from the above observations. To further understand the differing dynamic behaviour of water and IL observed around DNA base pairs, we have calculated the number of water molecules within 1 nm radius from different sites of DNA. The values are shown in Table 4.7. Interestingly, we find that the number of water molecules is always higher near CG-rich DNA, indicating a more crowded environment around CG base pairs. This crowding was further verified by calculating the average volume of the simulation box. We find that, although the number of water and IL

molecules are same for AT and CG-DNA systems, the box size is smaller for CG-DNA, reflected in the augmented density around CG compared to AT-rich DNA in IL solution.

Table 4.7 The number of water near different DNA sites and the simulation box volumes for the two dodecamer systems in different IL solutions.

IL solution	Water near major groove		Water near minor groove		Water near phosphate groups		Box volume (nm^3)	
	AATT	CGCG	AATT	CGCG	AATT	CGCG	AATT	CGCG
[BIMIM][Cl]	131	174	107	164	282	357	424	354
[Choline][Cl]	127	158	106	151	284	357	415	351
[BPYR][Cl]	134	176	117	168	283	364	420	350
[BPYR][PF_6]	154	212	133	204	329	446	361	299
[BPYR][BF_4]	140	201	122	188	296	409	391	331

Above results indicate the possible origin of slower dynamics of water and IL cations near GC base pairs. Due to higher density and more crowding, IL cations also get trapped near CG base pairs. Therefore, the slower dynamics of the water also slows down the dynamics of the IL cations. On the other hand, water molecules near AT base pairs are faster. Therefore, IL cations face lesser obstruction and show faster dynamics near AT base pairs despite having stronger interaction. Therefore, the origin of the dynamical behaviour of IL cations around DNA does not lie from the interaction with DNA; the dynamical behaviour of water modulates the dynamics of IL cations.

4.4 Discussion

The correlation between relaxation times of glassy systems and the configurational entropy (S_{conf}) has been known for a long time[43-44]. A liquid with slower dynamics may sample fewer configurations compared to a liquid with faster dynamics since some of the configurations will be inaccessible for a slow moving molecule. Consequently, the slower liquid will have lower value of S_{conf}. The diffusivity and the configurational entropy have been shown to have close correspondence at different temperature ranges, which is known as Adam-Gibbs relation that connects the dynamics of the medium with its entropy.[45],[46] Here, we bring this connection to

indicate how the difference in dynamics can impact solvation shell thermodynamics of the ILs. In the present study, we find the dynamics of IL cations to be significantly slower near CG-rich DNA. Therefore, the slower cations near CG rich segments could be in an entropically unfavourable state compared to those near AT base pairs. Moreover, the cations near CG base pairs remain in a denser environment. This will also make fewer configurations available for an IL cation near CG base pair compared to IL cation near AT base pairs. Previously, the minor groove of DNA with longer MRT values for water has been quantitatively shown to have lower entropy of water compared to the major groove.[47]We believe that the notable difference in the dynamics of IL and the different density environment of water molecules near AT and CG base pairs will contribute to the thermodynamics of IL cations binding with DNA. The IL cations near CG base pairs not only possess weaker enthalpic interaction due to their poor affinity towards IL cations, they also have lower entropic stabilization, resulting in an overall weak binding free energy. AT base pairs, on the other hand have the advantage of both enthalpy and entropic contributions for binding with IL cations.

Since water causes the slowdown of ionic liquids around CG, and conversely enhance the dynamics around AT, at higher concentration of IL, the effect of water will be smaller. Since ILs are known to have slower diffusion, at a very low water concentration, the dynamics near both AT and CG base pairs will be very slow. Therefore, we suggest that the entropic destabilization of CG-rich DNA will be insignificant and only the interaction energy of IL and DNA will possibly govern the stability. We could not probe the dynamics at low water concentrations due to extremely slow dynamics for which a significantly longer simulation and analysis will be required.

4.5 Conclusion

We have probed the dynamical behaviour of IL cations and water in hydrated IL systems near DNA with different AT and CG contents. We employed the stable state picture (SSP) approach for calculating the MRT values near major and minor groove and phosphate group atoms. We find that for the five different ILs considered in the study, mean residence time is always longer near CG-base pair sites compared to AT-base pairs. Similar behaviour was observed for water molecules also with MRTs near CG-base pair being higher compared to AT-base pairs, however, for different reasons. Our previous study[25] showed the MRT values in aqueous solution near different sites of DNA is mainly influenced by the position of the base pair with little dependence on the type of base step. Contrary to that finding, here the MRT values

obtained are significantly different for different kinds of base pairs kept at the same position. This observation indicates that in the hydrated IL solutions, chemical nature of different base pairs have a more significant role in altering the dynamics of the solvents. Our calculations indicated that although the IL cations interact weakly with CG base pairs compared to AT, water has much stronger interaction with CG-base pairs in the presence of ILs. Also, the reduction in simulation box volume for CG-DNA indicated the presence of a denser and crowded environment for this DNA. The pattern was found similar at different IL concentrations as well as in pure water. Due to the stronger interaction with CG base pairs, water has slower dynamics. This slow dynamics also influences the dynamics of IL. The slower dynamics also indicate lower entropy. Therefore, we have argued that both enthalpy and entropy disfavour binding of IL cations to CG and favour binding to AT. However, at a high concentration, enthalpy probably dominates the binding. The exact extent of entropic contribution in the stability of DNA can only be measured by the quantitative measurement, which will involve a separate rigorous future study. Overall our findings points to the fact that the measurements of dynamical quantities should also be considered to determine the efficacies of a new IL-based material for DNA storage.

4.6 References

1. Kutzler, M. A.; Weiner, D. B. *Nat. Rev. Genet.* **2008,** *9* (10), 776-788.
2. Phadke, R. S. *Appl. Biochem. Biotechnol.* **2001,** *96* (1-3), 269-276.
3. Yurke, B.; Turberfield, A. J.; Mills, A. P.; Simmel, F. C.; Neumann, J. L. *Nature* **2000,** *406* (6796), 605-608.
4. Boersma, A. J.; Megens, R. P.; Feringa, B. L.; Roelfes, G. *Chem. Soc. Rev.* **2010,** *39* (6), 2083-2092.
5. Park, S.; Sugiyama, H. *Angewandte Chemie International Edition* **2010,** *49* (23), 3870-3878.
6. Lindahl, T.; Nyberg, B. *Biochemistry* **1972,** *11* (19), 3610-3618.
7. Huddleston, J. G.; Visser, A. E.; Reichert, W. M.; Willauer, H. D.; Broker, G. A.; Rogers, R. D. *Green Chem.* **2001,** *3* (4), 156-164.
8. He, Y.; Li, Z.; Simone, P.; Lodge, T. P. *J. Am. Chem. Soc.* **2006,** *128* (8), 2745-2750.
9. Vijayaraghavan, R.; Izgorodin, A.; Ganesh, V.; Surianarayanan, M.; MacFarlane, D. R. *Angewandte Chemie International Edition* **2010,** *49* (9), 1631-1633.
10. Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990,** *112* (5), 2008-2010.

11. Parker, T. M.; Hohenstein, E. G.; Parrish, R. M.; Hud, N. V.; Sherrill, C. D. *J. Am. Chem. Soc.* **2013,** *135* (4), 1306-1316.

12. Rees, W. A.; Yager, T. D.; Korte, J.; Von Hippel, P. H. *Biochemistry* **1993,** *32* (1), 137-144.

13. Melchior, W. B.; Von Hippel, P. H. *Proc. Natl. Acad. Sci. U. S. A.* **1973,** *70* (2), 298-302.

14. Tateishi-Karimata, H.; Sugimoto, N. *Angew. Chem. Int. Ed.* **2012,** *51* (6), 1416-1419.

15. Ding, Y.; Zhang, L.; Xie, J.; Guo, R. *J. Phys. Chem. B* **2010,** *114* (5), 2033-2043.

16. Chandran, A.; Ghoshdastidar, D.; Senapati, S. *J. Am. Chem. Soc.* **2012,** *134* (50), 20330-20339.

17. Portella, G.; Germann, M. W.; Hud, N. V.; Orozco, M. *J. Am. Chem. Soc.* **2014,** *136* (8), 3075-3086.

18. Nakano, M.; Tateishi-Karimata, H.; Tanaka, S.; Sugimoto, N. *J. Phys. Chem. B* **2014,** *118* (2), 379-389.

19. Jumbri, K.; Abdul Rahman, M. B.; Abdulmalek, E.; Ahmad, H.; Micaelo, N. M. *Phys. Chem. Chem. Phys.* **2014,** *16* (27), 14036-14046.

20. Nakano, M.; Tateishi-Karimata, H.; Tanaka, S.; Sugimoto, N. *J. Phys. Chem. B* **2014,** *118* (32), 9583-9594.

21. Marušič, M.; Tateishi-Karimata, H.; Sugimoto, N.; Plavec, J. *Biochimie* **2015,** *108*, 169-177.

22. Kowsari, M. H.; Alavi, S.; Ashrafizaadeh, M.; Najafi, B. *J. Chem. Phys.* **2008,** *129* (22), 224508.

23. Bagchi, B. *Chem. Rev.* **2005,** *105* (9), 3197-3219.

24. Northrup, S. H.; Hynes, J. T. *J. Chem. Phys.* **1980,** *73* (6), 2700-2714.

25. Saha, D.; Supekar, S.; Mukherjee, A. *J. Phys. Chem. B* **2015,** *119* (34), 11371-11381.

26. Fujita, K.; MacFarlane, D. R.; Forsyth, M. *Chem. Commun.* **2005,** (38), 4804-4806.

27. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005,** *26* (16), 1668-1688.

28. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

29. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham Iii, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007,** *92* (11), 3817-3829.

30. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008,** *4* (3), 435-447.

31. Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005,** *88* (4), 2472-2493.

32. Sambasivarao, S. V.; Acevedo, O. *J. Chem. Theory Comput.* **2009,** *5* (4), 1038-1050.

33. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

34. Schlick, T. *Molecular modelling and simulation: An interdisplinary guide*. 2010; Vol. 2nd Ed.

35. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

36. Nose, S. *Mol. Phys.* **1984,** *52* (2), 255-268.

37. Hoover, W. G. *Phys. Rev. A* **1985,** *31* (3), 1695-1697.

38. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

39. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

40. Laage, D.; Hynes, J. T. *The Journal of Physical Chemistry B* **2008,** *112* (26), 7697-7701.

41. Tsuzuki, S.; Shinoda, W.; Saito, H.; Mikami, M.; Tokuda, H.; Watanabe, M. *J. Phys. Chem. B* **2009,** *113* (31), 10641-10649.

42. Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009,** *37* (17), 5917-5929.

43. Gibbs, J. H.; DiMarzio, E. A. *J. Chem. Phys.* **1958,** *28* (3), 373-383.

44. Adam, G.; Gibbs, J. H. *J. Chem. Phys.* **1965,** *43* (1), 139-146.

45. Scala, A.; Starr, F. W.; La Nave, E.; Sciortino, F.; Stanley, H. E. *Nature* **2000,** *406* (6792), 166-169.

46. Debenedetti, P. G.; Stillinger, F. H. *Nature* **2001,** *410* (6825), 259-267.

47. Jana, B.; Pal, S.; Maiti, P. K.; Lin, S.-T.; Hynes, J. T.; Bagchi, B. *J. Phys. Chem. B* **2006,** *110* (39), 19611-19618.

Chapter 5

Effect of Ions on Individual Water Entropy

5.1 Introduction

The studies discussed so far in the thesis have focused on the dynamics of water in the solvation shell. The different dynamic and structural aspects of water are often responsible for the thermodynamic properties of water, which will be the focus of this chapter.

Water in the environment is rarely found in absolutely pure form. The presence of different solutes affects the structure and properties of water in its vicinity differently[1]. For example, the presence of various ions makes the sea water participate in corrosion processes[2], whereas the ions in cellular water affect the stability of nucleic acids[3] and proteins[4]. Naturally, the hydration shell of ions has been subjected to numerous experimental and theoretical studies, especially due to the unique properties exhibited by various ions. The structure maker and breaker properties of ions[5] have been manifested in the viscosity of different ionic solutions. [6] Different ions also have been found to have different propensity to precipitate proteins in aqueous solutions leading to the Hofmeister's empirical classification. [3-4] The various properties are often considered to originate from the manner in which ions modulate the thermodynamics of its hydration water. Here we focus our study on the entropy of solvation shell water around ions, one at a time.

Entropy of water has been shown to be important in biologically relevant processes such as molecular recognition[7-8] and hydrophobic interactions[9-10].Therefore, several methods have been proposed for evaluating the entropic contribution of water from solvation dynamics. As discussed in Section 1.2.2 in Chapter 1, some of the tools used for evaluating thermodynamic properties of hydration water such as WaterMap[11-12], STOW[13], GIST[14-15] implement Inhomogeneous Solvation Theory (IST) [16-19], where solute-water pair correlation is used to estimate solvation entropy. Other methods such as 3D-RISM[20], SZMAP[21], GCT[22], 2PT decomposition method[23], etc. employ different statistical mechanical approaches. In spite of these developments, the systematic use of these methods has been hampered due to limited accuracy in some cases, the reasons for which has been discussed by Velez-Vega, *et al*.[24]The other method have used Potential Distribution Theorem (PDT) to obtain the excess chemical potential from which, hydration enthalpy and entropy can be calculated.[25] Most of these

methods are not designed to probe the water molecule at single particle level at a particular distance from the solute.

Here we apply the method for calculating individual water entropy, previously reported by our group[26] on the hydration shell of ions. Although we showed in the previous study[26] that our method captures the right trend in the hydrophobic crossover and a quantitative similarity with solvation entropy around methane, we did not perform an exhaustive comparison with experiment on other systems. We chose ionic solutions because of their variability in terms of size, charge, and interaction, while being simple enough to be tractable computationally. In addition, several aspects of ionic solutions are believed to originate from how the solvation shell water is modulated by ions. The structure maker tag for highly charged ions has been given to those which cause strong electrostatic ordering of nearby water molecules, reducing H-bonds among water molecules. [27] The extent to which these structural modifications happen in the solvation shell have inspired several studies, [28-31] none, however, with the entropic aspects of water. Dynamical aspects of water around ions also generated huge interest. Stirnemann, *et al.*[32] showed and subsequently explained using an extended jump model that while some ions tend to slow down the water near it, some other ions accelerate water next to it. The above studies thus provokes to ask how different ions modulate the entropic behavior of ions, which will also enlighten about the possible entropy behavior of water around various charged amino acids that act as a target in drug design studies.

In this study, we have calculated translational and rotational entropy of individual water molecules as a function of distance from different monovalent and divalent cations and anions. The accumulated change in the total entropy of water around an ion has been compared with experiment to establish the reliability of the method for both polarizable and non-polarizable force fields. We also measure the entropy change contributions from different solvation shells. Our study reveals distinct manner in which cations and anions modulate water entropy. The entropic behavior has been compared with structural properties of water around ions to establish the extent of impact ions have on properties of water. The results have been compared to several phenomena observed for water in case of ionic solutions and also for charged residues in water.

5.2 Methods and Simulation Details

5.2.1 System Set-up

For the calculation of entropy and structural properties of solvation shell water, we have used infinitely dilute ion concentration where one cation and one anion were solvated by 2000 water molecules. A distance restrain was put on the cation and anion so that ion pairs cannot form allowing us to calculate the water entropy in the solvation shell of individual ions independently. For cations, we have taken Li^+, Na^+, K^+, Rb^+, Cs^+, Mg^{2+}, Ca^{2+} and Zn^{2+} with keeping the anion as Cl^- for all the systems. For anions, we have taken F^-, Cl^-, Br^-, I^-, ClO^{4-} and SO_4^{2-} with the cation as Na^+ for all these systems. We have used polarizable models based on classical Drude oscillators for ions reported by Yu, *et al.*[33] except for ClO^{4-} and SO_4^{2-} whose parameters were obtained from Refs. [34] and [35], respectively. Polarizable SWM4-NDP water mode[36] was used for better depiction of the overall systems. We have compared the result with non-polarizable models of a few ions (Na^+, Cs^+, Mg^{2+}, Cl^-, I^-) with Amber99 force field[37-38] solvated by TIP4P[39] water model.

5.2.2 Simulation Details

All the simulations were carried out using GROMACS[40] software package. The ion-water systems were taken in a cubic box. Energy minimization were carried out using steepest descent method[41] followed by heating up to 298 K using Berendsen thermostat[42] with a coupling constant of 0.2 ps. The final equilibration was carried out for 2 ns at a constant temperature of 298 K and constant pressure of 1 bar using Nose-Hoover thermostat[43-44] and Parrinello-Rahman barostat[45], respectively with a coupling constant of 0.2 ps for each. Electrostatic interactions were treated using PME electrostatics[46] with 10 Å cut-off. Van der Waals cut-off was set to 10 Å. Choosing the average volume from the NPT equilibration mentioned above, a final molecular dynamics run under NVT condition for 50 ns was carried out for all the ion-water systems with all other parameters similar to the equilibration step. Distance restraint between the cation and anion was maintained throughout all the steps.

5.2.3 Method for Entropy and H-bond calculation

The method for single water translational and rotational entropy calculation has been discussed in Section 2.3 in Chapter 2. For studying the structural properties of water, the

average H-bonds per water have been calculated for permuted water molecules. Using the widely accepted convention[47] for H-bonding, two water molecules are considered to be forming H-bond if the distance between donor and acceptor is less than 3.5 Å and the donor-hydrogen-acceptor angle is less than 30°. For the H-bonding behavior around cations, only water-water H-bonding has been considered and for the case of anions, we have considered the interaction of ion with the hydrogen of water molecules in addition to the water-water H-bonding.

5.3 Results and Discussions

We begin our discussion with the translational and rotational entropy of individual water molecules in the solvation shell of ions. We will come back to the discussion of different solvation shells on the effect of entropy and water structure, thereafter.

5.3.1 Entropy Variation with Distance from the Ions

The translational ($T\Delta S_{trans.}$) and rotational entropy ($T\Delta S_{rot.}$) of individual water molecules around various ions is shown in Fig. 5.1 as a function of its distance from the respective ion. The 50 ns simulation employed here provides converged entropy values as shown in Fig. 5.2 for few water molecules around Cl^- as a representative ion. Entropy variation around water as a solute in a pure water system is shown in Fig. 5.3 for comparison. In Fig. 5.3, all the water molecules beyond 4 Å from the central water molecule are found, as expected, to be very similar. The average entropy for all these water molecules is taken to be the bulk water entropy. For the polarizable model, translational bulk entropy value is found to be 5.67 kcal/mol, while rotational bulk entropy is estimated to be -0.035 kcal/mol. Entropy estimates presented here are multiplied by 298 K to get the unit of free energy. The bulk entropy is subtracted from all the absolute entropy obtained for all other systems shown in Fig. 5.1.

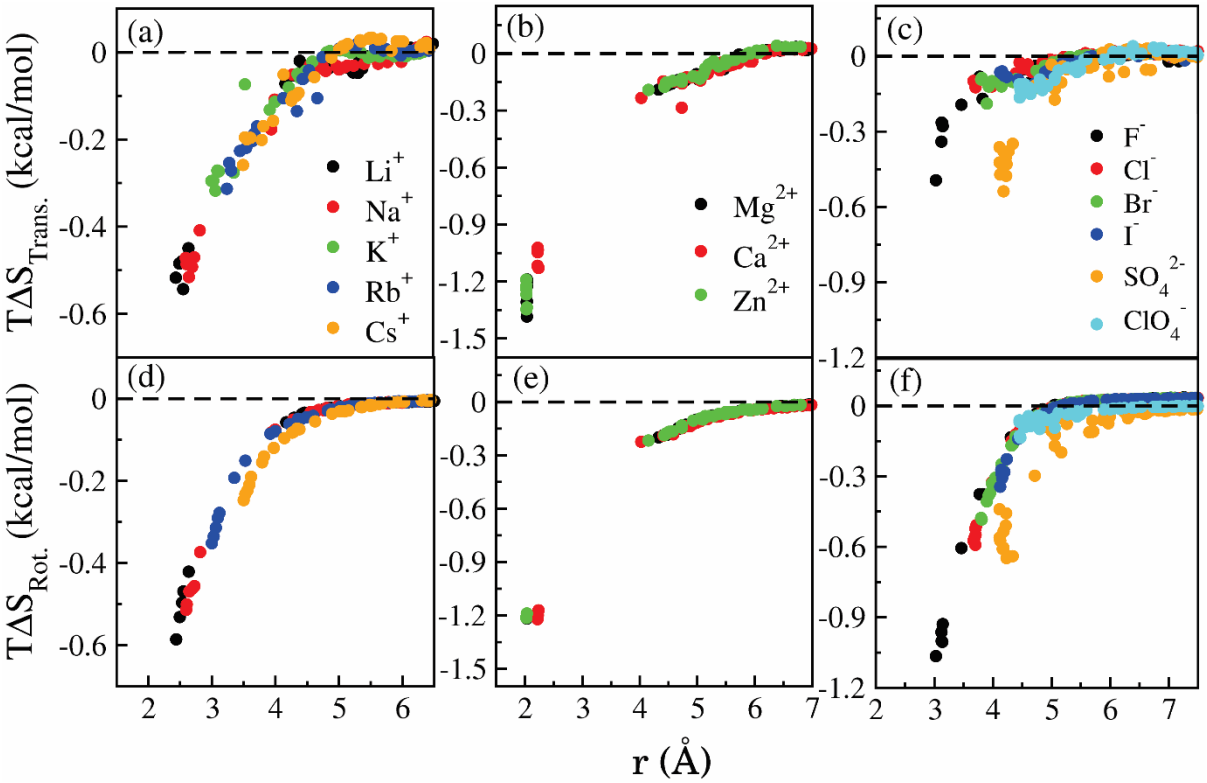

Figure 5.1 The deviation in translational entropy (top panel) and rotational entropy (bottom panel) from bulk value for monovalent cations (a and d), divalent cations (b and e), and anions (c and f). Entropy change is multiplied by the temperature 298K. The color code used in top and bottom panel represents same ions. Note, that scales of abscissa are same between top and bottom panels for each figure while ordinates are different in different figures.

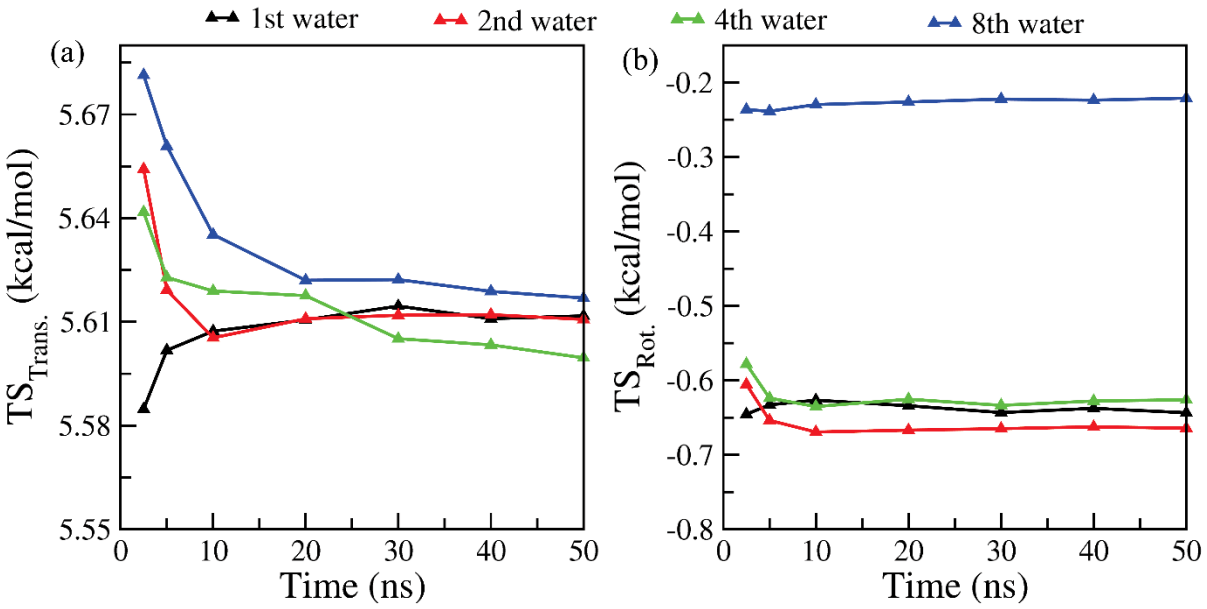

Figure 5.2 Convergence plot for (a) translational entropy and (b) rotational entropy value for four different water molecules around Cl^- ion.

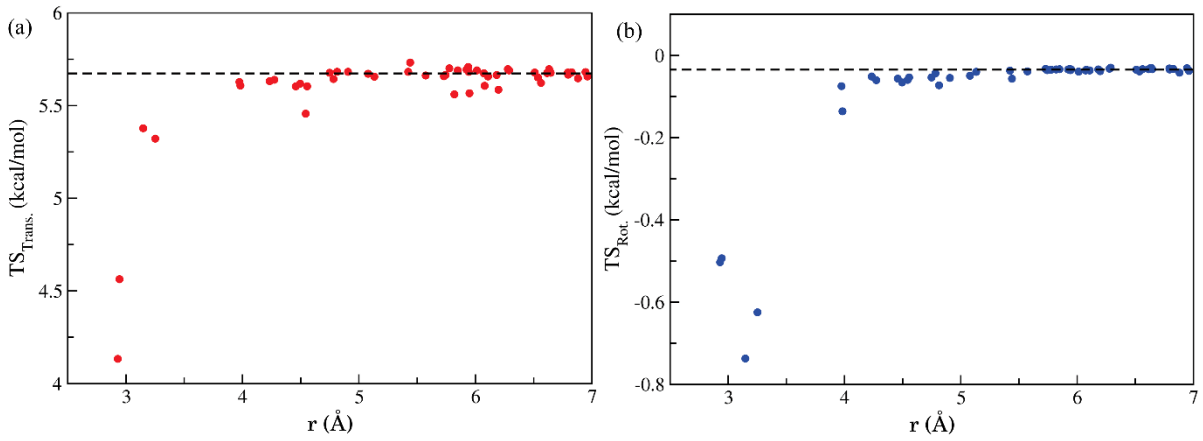

Figure 5.3 Variation of (a) translational entropy and (b) rotational entropy with distance for water around a certain water molecule. The bulk water value has been taken to be the average value at longer distance which has been shown by dashed line.

For all systems, the entropy of the water closest to the ion is the lowest. With increasing distance, entropy increases. Beyond a certain distance, entropy deviation from bulk value reaches zero, as expected. This result allows us not only to quantitatively predict the range within which ions impact water entropy; it shows the strength of the impact and number of water molecules that gets affected, as well. Figure 5.1(a) compares of the entropy variation for monovalent cations of different sizes, while Fig. 5.1(b) and 5.1(c) do the same for divalent cations and anions, respectively. Water molecules around the smallest cation Li^+ have the lowest entropy, followed by an increase in entropy for Na^+, K^+, Rb^+, and Cs^+, respectively, according to the increase in size. For Cs^+, which has a large hydrophobic surface affinity[48], a slight increase than bulk entropy is observed before the entropy reaches the bulk value. Closest set of water molecules around divalent cations loses more entropy than that around the monovalent cations. The increase in entropy with distance is steeper for divalent cations than the monovalent ones. Translational entropy behaves similarly for monovalent cations and anions. It is interesting to observe that anions decrease more rotational entropy than cations (compare F^- and Li^+). The divalent anions studied here are much bigger in size and induce a similar decrease in translational and rotational entropy as monovalent cations.

Before we go to further discussion, we need to show how reliable our estimations are. The entropy loss of individual water molecules around a particular ion is currently beyond the reach of experimental methods. Hence, we have estimated the total change in the water entropy from the translational and rotational entropy contributions of individual water molecules around each ion. Figure 5.4 shows the comparison between the total entropy change calculated as mentioned above with experimentally reported values[49] for both polarizable and non-polarizable force

fields. From the figure, we see that our calculated values correlate extremely well (correlation coefficient 0.97 and 0.98 for the two types of force fields respectively) with the experimental values. We also find our calculated values to be very similar to the entropic contribution to the hydration free energy of different ions calculated by Floriánet, *et al.*[50] using Langevin dipole model. While the comparison validates our method and result, it also indicates that the origin of the total entropy change is due to varying contributions of water molecules lying at different separations from the ions. The agreement for both polarizable and non-polarizable force field thus bolsters confidence on the methods adopted by us for entropy calculation and asserts the applicability of the method for various models and force fields.

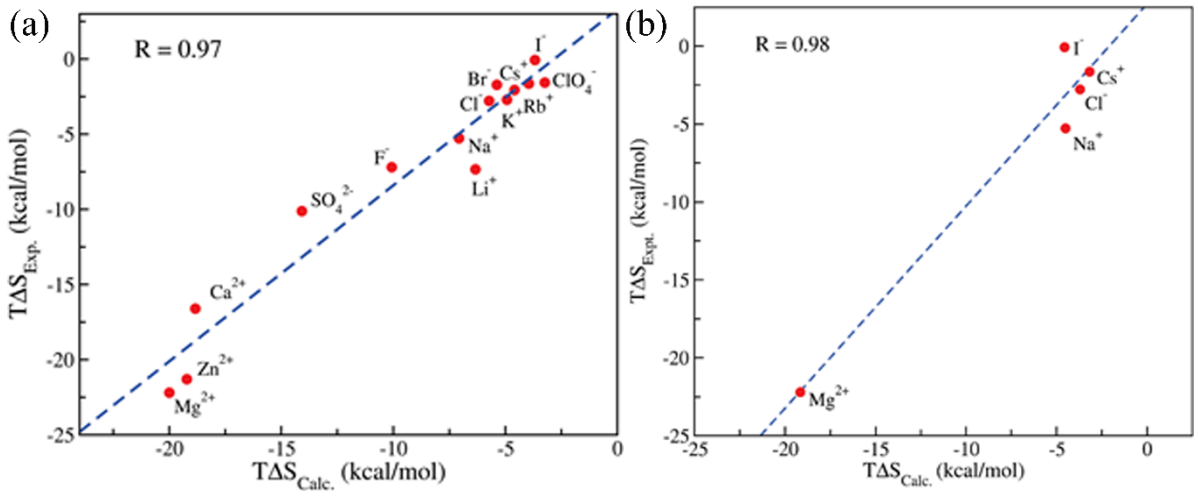

Figure 5.4 Comparison of experimental[49] and calculated entropy values for different ions at 298K. R is the Pearson's correlation coefficient.

The excellent correlation with previous reports implies that we can now obtain experimental values of the individual water entropy with just a constant difference from the theoretically calculated values. The method provides us with a powerful tool to have access of thermodynamics at the molecular level and as a function of distance from a solute.

5.3.2 Water Entropy in Different Solvation Shells

Since we have established the accuracy of our method, we can now discuss the entropy changes in different solvation shells of various ions. For this, we have calculated the radial distribution function (RDF) of water oxygen atoms around different ions and show the entropy change for individual water in Fig. 5.5 along with the total (translational + rotational) entropy change for all systems.

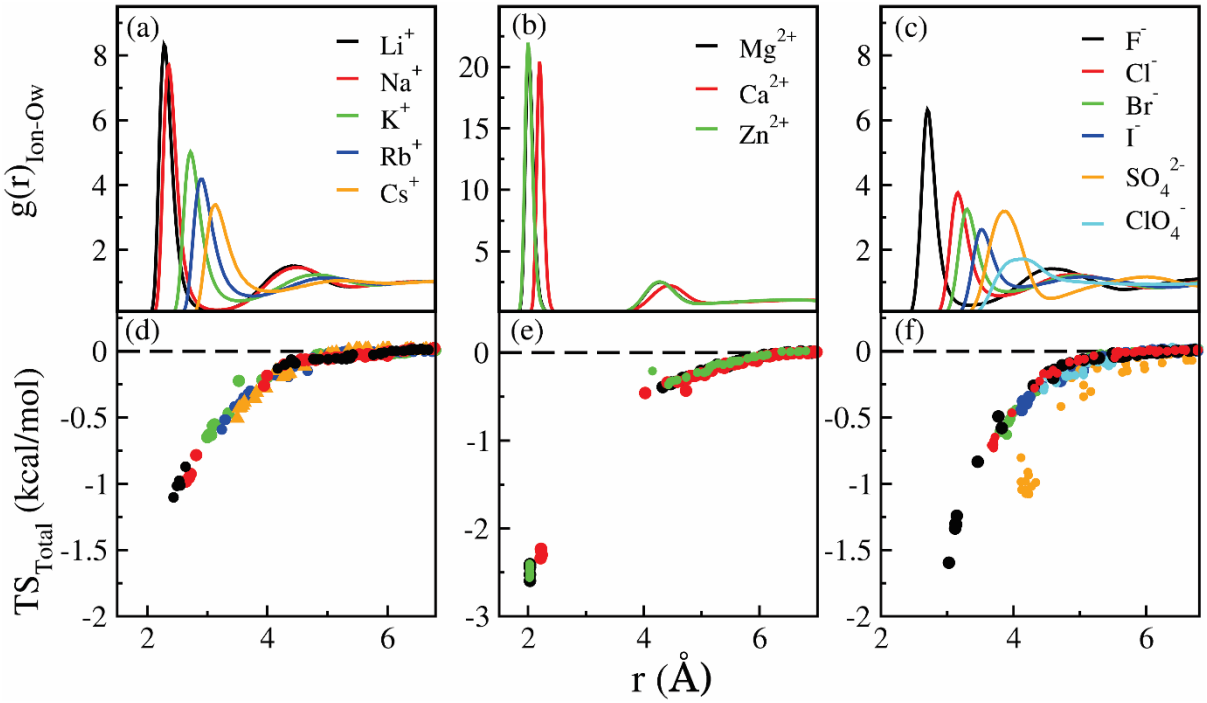

Figure 5.5 (Upper panel) Radial distribution function (RDF) plots between (a) monovalent cations-water oxygen, (b) divalent cations-water oxygen, (c) anions-water oxygen. (Lower panel) Single water entropy values for (d) monovalent cations, (e) divalent cations and (e) anions as a function of distance from the ion. Rest of the detail is same as Fig. 5.1.

Conventionally, RDF defines the regions of the solvation shells (first peak corresponding to the first solvation shell, second peak for the second solvation shell, etc.). The present approach, however, allows us to identify average positions of each molecule belonging within and even between the solvation shells. Note that for highly charged ions, there is a distinction between the first solvation shell and the second as can be seen from Fig. 5.5(e) where we see a gap between the closest six water and the rest. On the other hand, for ions larger than K^+ and anions in general, the solvation shells are more diffused due to their lower charge densities, and we see continuous presence of water with increasing entropy value. The pronounced second solvation shell was also observed around smaller singly charged and all doubly charged cations from the AIMD simulation study of Ding, *et al.*[51] The entropy reduction of the first solvation shell water molecules near monovalent cations is almost half compared to that around divalent cations. Therefore, both the size and the charge affect the entropy, indicating charge density shown in Table 5.1, to be a relevant parameter to compare the entropy reduction of water.

Table 5.1 The surface charge density for the spherical ions considered in this study.

Cation	Surface Charge density ($Å^{-2}$)	Anion	Surface Charge density ($Å^{-2}$)
Li^+	0.138	F^-	-0.045
Na^+	0.076	Cl^-	-0.024
K^+	0.042	Br^-	-0.021
Rb^+	0.034	I^-	-0.016
Cs^+	0.028		
Mg^{2+}	0.307		
Ca^{2+}	0.159		
Zn^{2+}	0.291		

However, charge density alone cannot explain the difference in the effect on water by cations and anions. Water molecules in the first solvation shell of F^- loose more entropy compared to Na^+ or even Li^+ that have higher charge density than F^-. The reason may lie in the capacity of the anions to form H-bond with water and thus have a greater effect in the rotational entropy as mentioned above. The entropy of individual water near halide ions correlates with the dynamical pattern of halide-water hydrogen bonds. H-bond is found to be strongest between water and F^- followed by Cl^-, Br^-, and I^- with all of them having a slower H-bond dynamics compared to the water-water H-bond.[52] Our calculations also show lower entropy values around these halides ions compared to that around bulk water.

Since the RDF gives the knowledge of different solvation shells, we can now probe the contribution of different solvation shells towards the total entropy change of water around a particular solute. Hence, the number of water in different solvation shells and their contributions towards the total entropy change in first two solvation shells are shown in Table 5.2. From the table, we see that the major contribution to the total entropy reduction around a solute come from the water in the first hydration shell. This pattern is in agreement with results of Beck[25] where the local electrostatic contribution has been found to vary with ion size, whereas the far-field electrostatic contributions remained similar for different ions, indicating the major contribution of first solvation shell. The second hydration shell, however, also has approximately 20% contribution across all ions, albeit with a larger number of water molecules. Therefore, while an individual water molecule in the first solvation shell contributes to more than 10% to the total entropy loss, a water molecule in the second solvation shell contributes 1% or less.

Table 5.2 The number of water molecules in different solvation shells and their contributions to the total entropy change for different ions.

Ion	Number^{FHS}	Contribution^{FHS} (%)		Number^{SHS}	Contribution^{SHS} (%)	
		$T\Delta S_{trans}$	$T\Delta S_{rot}$		$T\Delta S_{trans}$	$T\Delta S_{rot}$
Li^+	5	39.1	39.5	16	10.8	7.1
Na^+	6	40.3	39.2	16	10.6	7
K^+	7	36.6	38.8	19	10.8	12.2
Rb^+	8	40.7	34.2	20	11.6	13
Cs^+	9	41.2	40.8	21	3.5	14.1
Mg^{2+}	6	38.1	36.4	14	10.6	10.7
Ca^{2+}	6	33.8	38.3	14	10.6	9.9
Zn^{2+}	6	39.6	37.7	13	8.6	10
F^-	6	18.2	57.3	19	8.5	16
Cl^-	6	11.4	61.4	24	8.4	18.8
Br^-	7	14.8	55.7	23	8.4	21.1
I^-	7	14	57.3	26	12.3	16.4
SO_4^{2-}	12	33.8	46.2	55	4.8	15.2
ClO_4^-	14	50.3	33.3	57	0.4	14.6

FHS: Number of water molecule and contribution to the total entropy change for the first hydration shell; SHS: Number of water molecule and contribution to the total entropy change for the second hydration shell. $T\Delta S_{trans}$ and $T\Delta S_{rot}$ denotes percentage contribution of translational and rotational entropy to the total entropy loss.

The table also shows that cations reduce both translational and rotational entropy almost equally across different solvation shells. The mechanism of entropy reduction around anions, on the contrary, stems mainly from a much greater reduction (>50% in the first solvation shell) of rotational entropy loss, indicating that charge density of ions may mainly affect the translational entropy more than the rotational entropy. Interestingly, residence time, a measure

of translational motion, of water around Li^+ is found to be much higher compared to F^-.[53] Therefore to establish the relation between charge density and different entropy contributions, we have plotted the translational and rotational entropy change of water in the first solvation shell of various ions against the magnitude of their charge densities. The plots are shown in Fig. 5.6. While translational entropy is highly correlated (correlation coefficient -0.95) with charge density, correlation of the same with rotational entropy is rather poor (correlation coefficient -0.78).

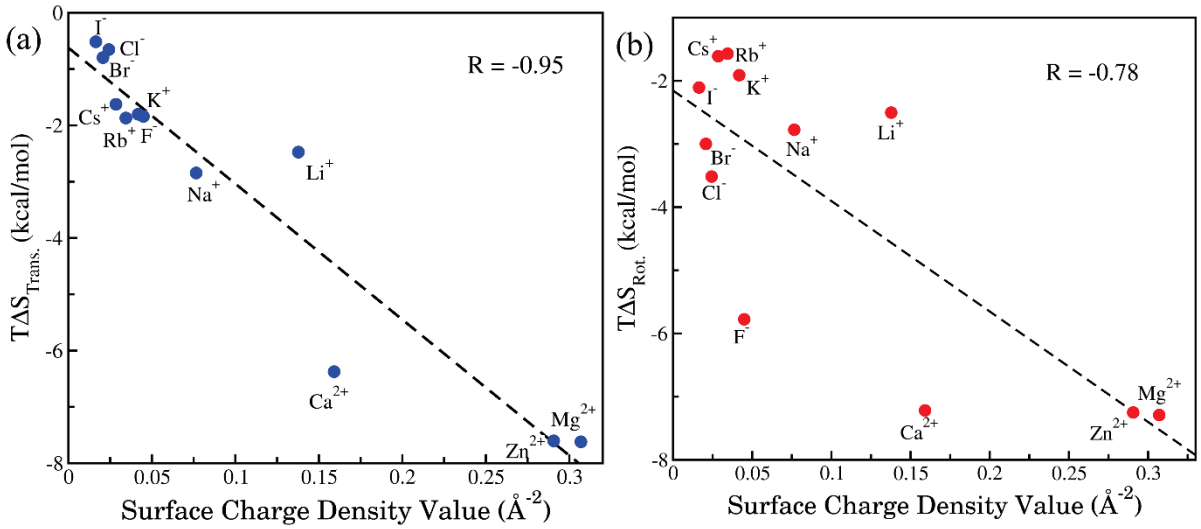

Figure 5.6 Correlation between translational entropy (a) and rotational entropy (b) with the magnitude of surface charge density values of cations and anions.

Therefore a clear trend between charge density and translational entropy can be observed from the above plot. But, anions affect rotational entropy more than the translational entropy. We probed into the origin of this behavior through examining water structure in the hydration shells.

5.3.3 Water Entropy and Water Structure Correlation

To investigate the origin of the entropic behavior, we have looked at the structural aspects of solvation water through its hydrogen bonding (H-bond) property. The average number of H-bond made by particular water at a certain distance was calculated for permuted water molecules. The values have been plotted with respect to the average distance of water molecules from the ions and are shown in Fig 5.7. Mostly for all the ions (except F^-), we observe an increase in the average H-bond with distance before reaching bulk value (dashed line) of 3.61, as it should, beyond a certain distance.

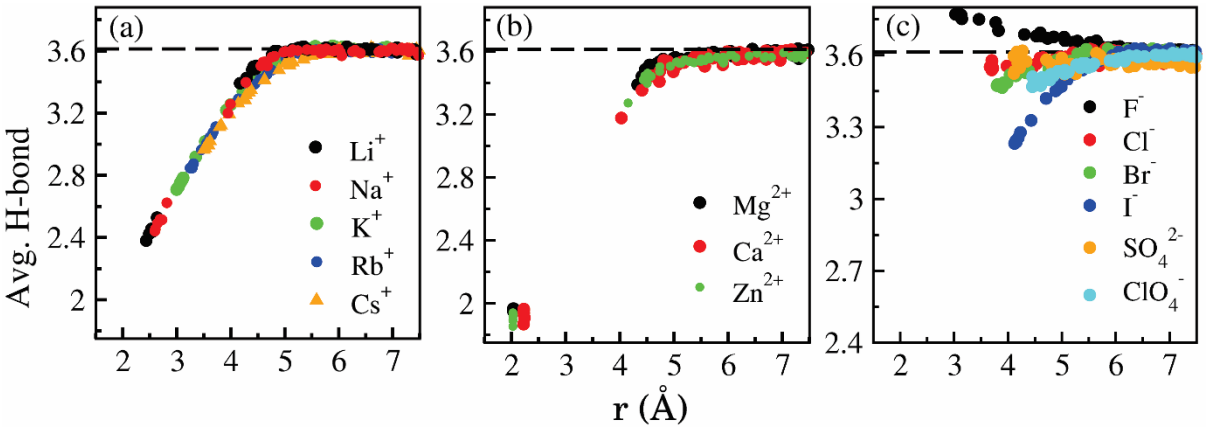

Figure 5.7 Average H-bonds per water molecule vs. distance of water plot for (a) monovalent cations, (b) divalent cations and (c) anions. The dashed lines represent the bulk water value. Note that ordinate range is different for different figures.

We find here that the average H-bond and entropy values of individual water molecules show a similar trend like entropy for cations: the water in the first shell has much lower H-bond and the ones in the second shell show a little decrease from bulk. H-bond values reach the bulk value at around the same distance as entropy, indicating that the extent of modulation of thermodynamic and structural properties by cations is similar. The anions, however, show a reverse trend, with F^- producing an enhanced H-bonding for the neighboring water molecules in spite of a high reduction in water entropy. The I^- ion with much smaller charge density show only a tiny reduction in entropy but a significant decrease in H-bonding. There may be two reasons for this reversal in behavior. Both cations and anions can hinder H-bond formation among water molecules by creating electrostatic ordering. While cations cannot compensate this loss in average H-bond by participating in H-bond formation, anions can accept H-bond from water molecules. While, I^- compensates this loss partially, F^- overcompensates it to have more H-bond than bulk.

The reduction of H-bonds reflects the structure maker property of the cations as argued by Hribar, *et al.*[27]. Since structure maker property and entropy are correlated with each other, a correlation between H-bond and entropy is also expected. Our results are in good agreement with experimental observations of Harsányi, *et al.*[54] that showed an increase in LiCl salt concentration breaks the H-bond network of water. The study by Mancinelli, *et al.*[29] on NaCl and KCl salt solutions using neutron diffraction/EPSR technique also found these solutions to behave in a manner similar to water at high pressure, indicating an enhanced ordering of water, especially for Na^+ ion compared to K^+. The pattern we see for H-bonding and entropy of

different ions on its surrounding water also causes several observable phenomena in water. Smaller effects on entropy by larger ions indicates a weaker interaction between these ions and water, as found in the case of Br^- and I^- by Karmakar, *et al.*[55-56]. The weaker interaction leads to a higher diffusion for the water near larger ions.[57] The lowering of entropy, on the other hand, is a related to stronger H-bonding. Interestingly, the different manifestations of structural (H-bond) and thermodynamic (entropy) properties presented here find a similar correlation in dynamical properties also where F^- and I^- show different modes of reorientation pathways.[58] Thus our study on individual water molecules accurately represents the previously reported results obtained for overall solvation shell of ions.

Different experimental studies provide water perturbation by ions to various extents. While the effects have been considered to prevail only up to the first solvation shell by some studies[28, 59], the work by Mancinelli, *et al.*[29] claimed the effects to be present even in the second and third solvation shells. Our results so far have shown that the largest impact on entropy and H-bonding behavior for individual water is restricted mainly to the first solvation shell, with some effect in the second shell. It is therefore interesting to see how different ions affect the orientations of individual water molecules present at different distances from the ion. It can also shed light on the rotational entropy pattern observed for anions. Ding, *et al.* also looked at the average orientational distribution of water molecules in the first solvation shell of NaCl and CsI using *ab initio* molecular dynamics simulations. [51] The difference is that here we can probe into the orientation of individual water molecules at various distances, not just the average of the solvation shell. We have calculated the angle θ between the vector from water oxygen to the ion and the vector bisecting the water molecule as shown in Fig. 5.8. Distributions of θ for four water molecules are plotted for several ions. Water around smaller cations (higher charge density) shows a peak around 166° for bivalent cations and decreases with decreasing charge density (Fig. 5.8a-d). The high θ value indicates that water hydrogens orient away from the cation, as expected. Water around larger cations has a broader distribution. This distribution is maintained for first six water molecules, i.e, the first solvation shell water. However, the trend vanishes from the seventh water onwards, and no pattern was found for remaining water molecules.

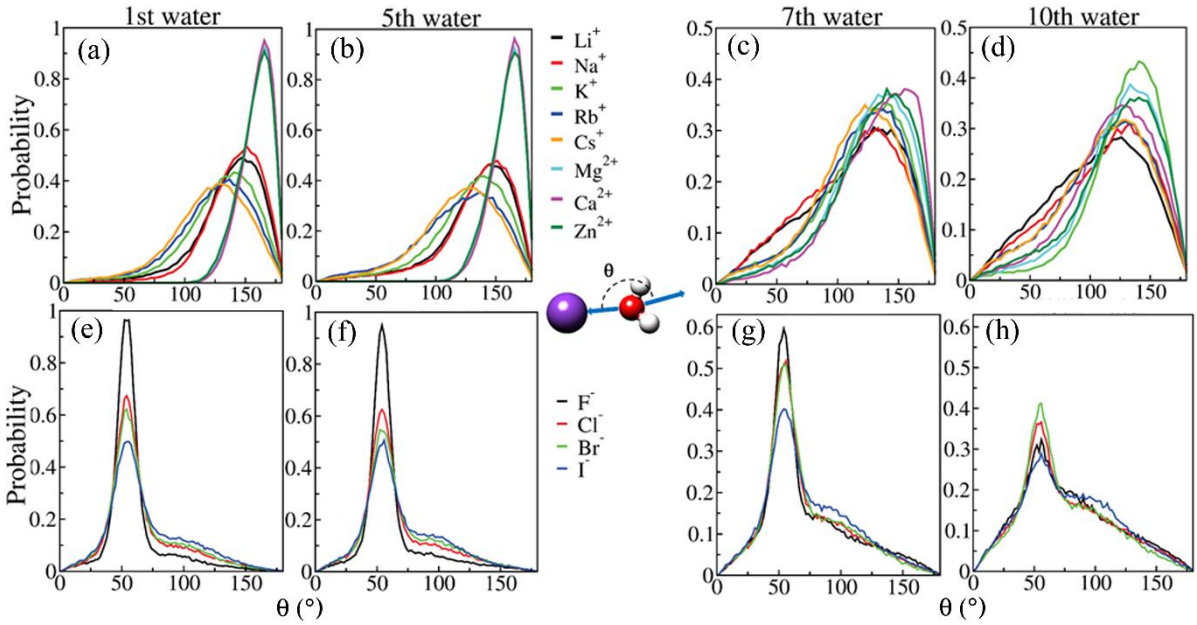

Figure 5.8 Distribution of the angle θ (shown in the middle) between the vector connecting the ion and the water oxygen atom with the bisector vector of the water molecule for cations (upper panel) and anions (lower panel). (a) and (e) plot is for first water, (b) and (f) plot is for fifth water, (c) and (g) plot is for seventh water and, (d) and (h) plot is for tenth water next to cations and anions respectively.

For anions (Fig. 5.8e-h), as expected, the peak occurs at smaller angle of 50° indicating that one of the hydrogen is H-bonded to the anion. For anions too, the highest peak is observed for F^- for up to seventh water, after which the orientation do not follow any pattern. Interestingly, for anions, even the 10th water shows a peak at around 50°, indicating a rather long-ranged effect compared to cations. This long-range effect might cause a greater loss of entropy of water around anions compared to cations of similar or even higher charge density. It also explains the higher rotational entropy loss in the second solvation shell for anions as shown in Table 5.2. Our previous study[26] also showed the entropy reduction for particles with -1 charge is higher compared to particle of the same size with +1 charge. We believe that this behavior of anions has implications towards the dynamics of water near biomolecules. Sterpone, *et al.*[60] calculated H-bond dynamics near different amino acids and reported H-bond axis tumbling time (τ_{reor}^{frame}). The study showed that the timescales near donor sites (positively charged residues) are smaller compared to the acceptor sites (negatively charged residues), indicating the water near these sites are held more strongly compared to donor sites. The dynamical features mentioned corroborate the thermodynamic findings of the present study where an acceptor groups like anions or negatively charge amino acid will have stronger impact on water than cations or positively charged amino acid that act as a donor.

5.4 Conclusion

In this part of the thesis, we report the entropy change of individual water molecules in different solvation shells of ions. The sum of contributions of individual water entropy around a particular ion correlates extremely well with experiment for both the polarizable and non-polarizable force fields. This study also reveals a fundamental difference in the approach of cations and ions by which they reduce entropy of the surrounding water molecules. Our quantitative measurements at individual water level provide the exact extent of the perturbations caused by ions on surrounding water molecules, albeit for lower concentration of ions. The similarity between entropic and H-bonding behavior show that the manner in which ions modulate thermodynamics of surrounding water have consequences towards the structure maker and breaker properties of these ions. The study reveals that charge density governs the translational entropy of solvent molecules, whereas the rotational entropy involves the strength of the interaction between ion and water. This manifests in differing roles of cations and anions on the water entropy. The general pattern found across all systems is that ions mostly affect the first solvation shell water. However, significant contribution (20%) to the entropy reduction comes from the second solvation shell. The scenario will be likely to change with increase in concentration where the solvation shells will overlap with different ions. Finally, our method successfully reproduces (within a constant) the experimental solvation entropy while retaining the information at the individual water level. The success of the present method implies that it can be applied to find the entropic behavior of water around proteins and DNA that govern important roles in drug design.[11, 61]

5.5 References

1. Marcus, Y.; Ben-Naim, A. *J. Chem. Phys.* **1985,** *83* (9), 4744-4759.
2. Kritzer, P. *J. Supercrit. Fluids* **2004,** *29* (1–2), 1-29.
3. Record Jr, M. T.; Zhang, W.; Anderson, C. F. *Adv. Protein Chem.* **1998,** *51*, 281.
4. Lo Nostro, P.; Ninham, B. W. *Chem. Rev.* **2012,** *112* (4), 2286-2322.
5. Marcus, Y. *J. Chem. Soc., Faraday Trans. 1* **1986,** *82* (1), 233-242.
6. Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945,** *13* (11), 507-532.
7. Baron, R.; Setny, P.; Andrew McCammon, J. *J. Am. Chem. Soc.* **2010,** *132* (34), 12091-12097.
8. Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. *Nature* **2007,** *448* (7151), 325-U3.

9. DeLorbe, J. E.; Clements, J. H.; Teresk, M. G.; Benfield, A. P.; Plake, H. R.; Millspaugh, L. E.; Martin, S. F. *J. Am. Chem. Soc.* **2009,** *131* (46), 16758-16770.

10. Smith, D. E.; Zhang, L.; Haymet, A. D. J. *J. Am. Chem. Soc.* **1992,** *114* (14), 5875-5876.

11. Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (3), 808-813.

12. Pearlstein, R. A.; Hu, Q. Y.; Zhou, J.; Yowe, D.; Levell, J.; Dale, B.; Kaushik, V. K.; Daniels, D.; Hanrahan, S.; Sherman, W.; Abel, R. *Proteins: Struct., Funct., Bioinf.* **2010,** *78* (12), 2571-2586.

13. Li, Z.; Lazaridis, T. *Methods Mol. Biol.* **2012,** *819*, 393-404.

14. Nguyen, C. N.; Young, T. K.; Gilson, M. K. *J. Chem. Phys.* **2012,** *137* (4).

15. Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. *J. Chem. Theory. Comput.* **2014,** *10* (7), 2769-2780.

16. Wallace, D. C. *J. Chem. Phys.* **1987,** *87* (4), 2282-2284.

17. Ashbaugh, H. S.; Paulaitis, M. E. *J. Phys. Chem.* **1996,** *100* (5), 1900-1913.

18. Lazaridis, T. *J. Phys. Chem. B* **1998,** *102* (18), 3542-3550.

19. Chang, C. E. A.; McLaughlin, W. A.; Baron, R.; Wang, W.; McCammon, J. A. *P. Natl. Acad. Sci. U. S. A.* **2008,** *105* (21), 7456-7461.

20. Sindhikara, D. J.; Hirata, F. *J. Phys. Chem. B* **2013,** *117* (22), 6718-6723.

21. Grant, J. A.; Pickup, B. T.; Nicholls, A. *J. Comput. Chem.* **2001,** *22* (6), 608-640.

22. Gerogiokas, G.; Calabro, G.; Henchman, R. H.; Southey, M. W. Y.; Law, R. J.; Michel, J. *J. Chem. Theory. Comput.* **2014,** *10* (1), 35-48.

23. Lin, S.-T.; Blanco, M.; Goddard, W. A. *J. Chem. Phys.* **2003,** *119* (22), 11792-11805.

24. Velez-Vega, C.; McKay, D. J. J.; Kurtzman, T.; Aravamuthan, V.; Pearlstein, R. A.; Duca, J. S. *J. Chem. Theory. Comput.* **2015,** *11* (11), 5090-5102.

25. Beck, T. L. *J. Phys. Chem. B* **2011,** *115* (32), 9776-9781.

26. Sasikala, W. D.; Mukherjee, A. *J. Phys. Chem. B* **2014,** *118* (36), 10553-10564.

27. Hribar, B.; Southall, N. T.; Vlachy, V.; Dill, K. A. *J. Am. Chem. Soc.* **2002,** *124* (41), 12302-12311.

28. Omta, A. W.; Kropman, M. F.; Woutersen, S.; Bakker, H. J. *Science* **2003,** *301* (5631), 347-349.

29. Mancinelli, R.; Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. *Phys. Chem. Chem. Phys.* **2007,** *9* (23), 2959-2967.

30. Paschek, D.; Ludwig, R. *Angew. Chem. Int. Edit.* **2011,** *50* (2), 352-353.

31. Tielrooij, K. J.; Garcia-Araez, N.; Bonn, M.; Bakker, H. J. *Science* **2010,** *328* (5981), 1006-1009.

32. Stirnemann, G.; Wernersson, E.; Jungwirth, P.; Laage, D. *J. Am. Chem. Soc.* **2013,** *135* (32), 11824-11831.

33. Yu, H. B.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D.; Roux, B. *J. Chem. Theory. Comput.* **2010,** *6* (3), 774-786.

34. Baer, M. D.; Kuo, I. F. W.; Bluhm, H.; Ghosal, S. *J. Phys. Chem. B* **2009,** *113* (48), 15843-15850.

35. Vila Verde, A.; Lipowsky, R. *J. Phys. Chem. B* **2013,** *117* (36), 10556-10566.

36. Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell Jr, A. D. *Chem. Phys. Lett.* **2006,** *418* (1–3), 245-249.

37. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

38. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007,** *92* (11), 3817-3829.

39. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996,** *118* (45), 11225-11236.

40. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory. Comput.* **2008,** *4* (3), 435-447.

41. Schlick, T. *Molecular modelling and simulation: An interdisplinary guide*. 2010; Vol. 2nd Ed.

42. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.

43. Nose, S. *Mol Phys* **1984,** *52* (2), 255-268.

44. Hoover, W. G. *Physical Review A* **1985,** *31* (3), 1695-1697.

45. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.

46. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.

47. Laage, D.; Hynes, J. T. *Science* **2006,** *311* (5762), 832.

48. Schwierz, N.; Horinek, D.; Netz, R. R. *Langmuir* **2013,** *29* (8), 2602-2614.

49. Marcus, Y. *Ion properties*. Marcel Dekker: New York, 1997.

50. Florián, J.; Warshel, A. *J. Phys. Chem. B* **1999,** *103* (46), 10282-10288.

51. Ding, Y.; Hassanali, A. A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111* (9), 3310-3315.

52. Chowdhuri, S.; Chandra, A. *J. Phys. Chem. B* **2006,** *110* (19), 9674-9680.

53. Lee, S. H.; Rasaiah, J. C. *J. Phys. Chem.* **1996,** *100* (4), 1420-1425.

54. Harsányi, I.; Pusztai, L. *J. Chem. Phys.* **2005,** *122* (12), 124512.

55. Karmakar, A.; Chandra, A. *Chem. Phys.* **2015,** *448*, 1-8.

56. Karmakar, A.; Chandra, A. *J. Phys. Chem. B* **2015,** *119* (27), 8561-8572.

57. Kim, J. S.; Wu, Z.; Morrow, A. R.; Yethiraj, A.; Yethiraj, A. *J. Phys. Chem. B* **2012,** *116* (39), 12007-12013.

58. Boisson, J.; Stirnemann, G.; Laage, D.; Hynes, J. T. *Phys. Chem. Chem. Phys.* **2011,** *13* (44), 19895-19901.

59. Omta, A. W.; Kropman, M. F.; Woutersen, S.; Bakker, H. J. *J. Chem. Phys.* **2003,** *119* (23), 12457-12461.

60. Sterpone, F.; Stirnemann, G.; Hynes, J. T.; Laage, D. *J. Phys. Chem. B* **2010,** *114* (5), 2083-2089.

61. Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008,** *130* (9), 2817-2831.

Chapter 6

Effect of Local Structure on the Thermodynamics of Individual Water at Supercooled State

6.1 Introduction

One of the reasons for water being able to keep researchers engaged over the years is its unique properties at different thermodynamic conditions. While the temperature dependence of different thermodynamic response functions such as isothermal compressibility, specific heat and thermal expansion coefficient show their distinctive patterns[1-3], the temperature dependence of water density can be viewed to be of critical importance with real life implications. The various anomalies in thermodynamic and kinetic properties of water play crucial roles in several chemical and physical processes in the fields of materials science, biology and geological processes.[4-7] A major breakthrough towards understanding these anomalies of water have emerged from the phase diagram of metastable water[8-9]. A number of attempts have been made to interpret the phase behavior of water with different theoretical scenarios.[7, 9-13] The most popular among these have been the concept of the existence of a liquid-liquid critical point (LLCP)[9] that marks the terminal point in the first order phase transition between high density liquid (HDL) and low density liquid (LDL) water.[14-16] While the existence of these two types of water at different conditions is widely accepted,[17-18] their co-existence at ambient condition has given rise to different opinions. While a number of studies claim that water at ambient condition remain as a mixture of HDL and LDL liquids with varying proportions at different temperatures,[19-21] others have claimed to find only one type of water at these conditions. [22-23] While the structural aspects of water have been proven to be crucial for explaining water anomalies[24], the true nature of water structure at ambient conditions and its consequence on the water thermodynamic properties remain ambiguous.

The other significant characteristic of water at supercooled state is its divergence of dynamic properties. The dynamic properties of water and other such glass forming liquids are reflected in their macroscopic properties such as viscosity or diffusion coefficient. Depending on the temperature dependence of these properties, these systems behave as strong or fragile liquids. This has been shown for several systems in the study by Mallamace *et al.*[25] The molecules at low temperatures in these systems become spatially correlated with different length scales.[26-28] As a consequence, the dynamics in water are dominated by rearrangement of

dynamically correlated regions through activated processes. Water being a glass forming liquid is known to undergo transition in its dynamic behavior at around 225 K.[29-30] This phenomenon is known as "fragile-to-strong" crossover (FSC) transition. The origin for such behavior for water comes from the ordering of water molecules as shown by Moore *et al.*[22]with an increase in four co-ordinated H-bonds at nearly FSC transition temperature. With enhanced structure for the water molecules, the dynamics are dominated by activated processes in water. Evidences for these processes at low temperatures have been investigated for both bulk water[31-32] and for water in the confined state[33] which have improved the understanding of this behavior. However, how the thermodynamic properties of water, particularly its entropy, are affected by water dynamic properties at various temperatures is not yet clear.

In this chapter, we intend to probe the consequences of water's above mentioned structural and dynamic features at different conditions on its translational ($TS_{Trans.}$) and rotational entropy ($TS_{Rot.}$) at single molecule level. We have employed a method discussed in Section 2.2.3 in Chapter 2 that calculates the single water entropy at a particular distance from the solute[34]. We have attempted to connect the translational diffusion of water with the $TS_{Trans.}$ values. Additionally we have investigated the individual water entropy values at low temperatures in more detail using a structural parameter that can distinguish differently structured water. With our approach of investigation at the single molecule level, we have found new insights on the structural polymorphism of water and its effect on water entropy. Using our findings, we attempt to bridge the different claims provided through numerous studies. Overall, our study aims to signify the importance of looking at individual water behavior at these conditions in order to address the inconsistencies found in previous reports by various groups in addition to adding new information to the features of supercool water.

6.2 Simulation Details

All atom molecular dynamics simulations were performed in our study to find the properties of water at different conditions. To accurately study water behavior at different temperatures, we have used TIP5P/E water model[35] which is known to give correct density pattern at different temperatures. The system was first energy minimized using steepest descent method[36]. This was followed by heating the systems up to 300 K using Berendsen thermostat[37] using a coupling constant of 0.2 ps. The system was then equilibrated for 2 ns at 300 K and 1 bar pressure using Nosé-Hoover thermostat[38-39] and Parrinello-Rahman barostat[40] with a

coupling constant of 0.4 ps for both. The equilibrated system thus obtained was then subjected to heating or cooling up to the temperatures at which the simulations were needed to be carried out. In this study, we have performed simulations at 190, 220, 230, 240, 250, 260, 273, 285, 300, 315 and 330 K. Therefore the systems were equilibrated at these temperatures using the above mentioned thermostat and barostat. The equilibration for systems below and at 273 K was carried out for 30 ns while systems above this temperature were subjected to 10 ns of equilibrations. A final simulation under NPT condition was carried out for each system with a time step of 2 fs. The trajectory frames were saved at every 0.1 ps from which the entropy values and other properties were calculated. These trajectories were 1000 ns long for systems at temperatures 190, 220 and 230 K while for systems simulated between 240 K and 273 K, the trajectories were 200 ns long. The systems simulated above these temperatures were simulated for 100 ns within which converged entropy values were obtained. The electrostatics were calculated with PME electrostatics[41] with a cut-off length of 10 Å for both electrostatics and van der Walls interactions. All the simulations were performed using GROMACS software package.[42] The method for entropy calculation has been discussed in Section 2.2.3 in Chapter 2. The reliability of this method has been discussed in Section 5.3.1 in Chapter 5. The diffusion coefficient of water has been calculated from the velocity auto-correlation function of water molecules using the following relation:

$$D = \frac{1}{d}\int_0^\infty \langle \vec{V}(0).\vec{V}(t)\rangle, \qquad (6.1)$$

where d is the dimensionality, D is the diffusion coefficient, and $\vec{V}$ is the velocity vector. The VACF profiles for water at different temperatures are shown in Fig. 6.1.

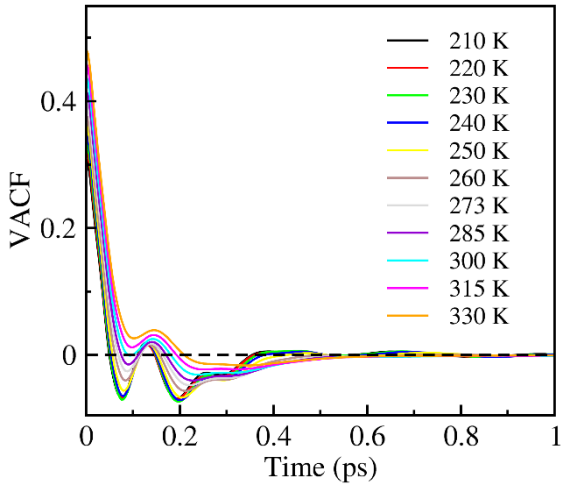

Figure 6.1 Time dependence of the velocity auto correlation function (VACF) of water at different temperatures.

6.3 Results

We first discuss effect of FSC transition on the diffusion behavior of water to show the effect of temperature change on water behavior. This is followed by investigation of how the diffusion is correlated with translational entropy of water. Then we look at the effect of FSC transition on translational and rotational entropy behavior at individual molecule level to see the temperature effects on single water entropy values. The structural aspect of water and their effect on water entropy behavior is then discussed in the subsequent sections.

6.3.1 Water Dynamics and Entropy at Different Temperatures:

We begin the discussions by looking at the diffusion coefficient of water at various temperatures. The logarithm of the inverse of diffusion coefficient ($\ln(D^{-1})$) has been plotted with respect to temperature in Fig. 6.2. From the figure, we can see that there are two regions present in the plot: the one at low temperature can be described using Arrhenius behavior, while the high temperature region can be fitted with the Vogel-Fulcher-Tammann (VFT) equation of the following form:[43]

$$D^{-1}(T) = D_0^{-1} \exp\left[\frac{1}{K_{VFT}\left(\frac{T}{T_{VFT}}-1\right)}\right], \qquad (6.2)$$

where K_{VFT} is the kinetic fragility marker and T_{VFT} is the temperature at which the diffusivity diverges. D_0 is the high temperature diffusivity. The two regions cross at around 250 K, which is known as the FSC transition temperature.[29] Here we find that the K_{VFT}, known as the dynamic fragility marker, to be 0.77 and T_{VFT}, the temperature at which glass transition takes place to be 204.8 K. These values are close to the ones reported by Chen et al.[44] where experiments showed through confinement of water in nanopores that the value of T_{VFT} and K_{VFT} are 187 K and 0.56, respectively.

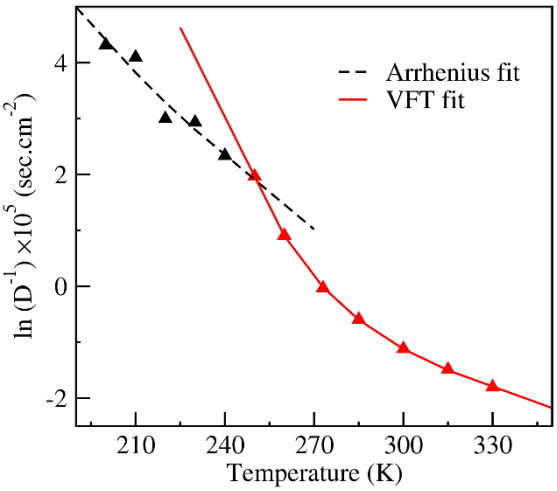

Figure 6.2 The logarithm of inverse diffusion coefficient with respect to temperature. The red solid line here shows the VFT fit, while the black dashed line shows the Arrhenius fit.

Now we look at the variation of translational entropy and rotational entropy, $TS_{Trans.}$ and TS_{Rot}, values with respect to temperatures. Note that, we have multiplied the temperature with the entropy to get the energy unit. Also, here the entropy values are associated with individual water molecules, albeit in the bulk water every water molecule should have same entropy. Therefore, for each temperature about 230 K, we have calculated the entropy of 30 individual water molecules and averaged the values. For systems at 190 and 220 K, due to a larger variation in entropy values for different water molecules, an average of 100 water molecules has been considered. Fig. 6.3 shows the $TS_{Trans.}$ and TS_{Rot} at different temperatures. The convergence of entropy values at 190 K are shown in Fig. 6.4. Simulations at higher temperature converges even faster.

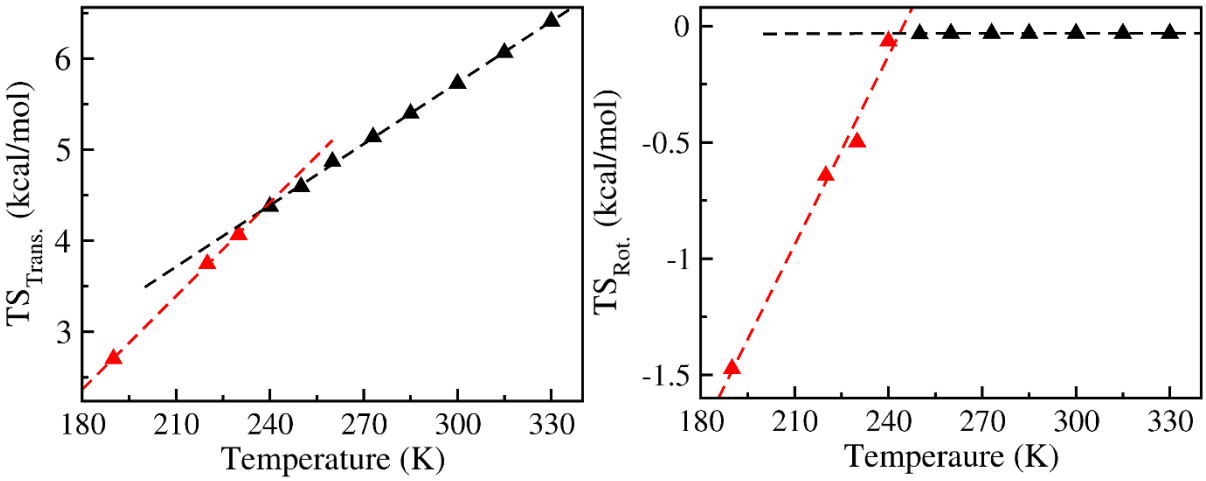

Figure 6.3 The average translational (TS_{Trans}; left) and rotational (TS_{Rot}; right) entropy of individual water molecules at different temperatures. The black and red points represent entropy values above and below the transition temperature, respectively. The dashed lines are used to guide the eyes.

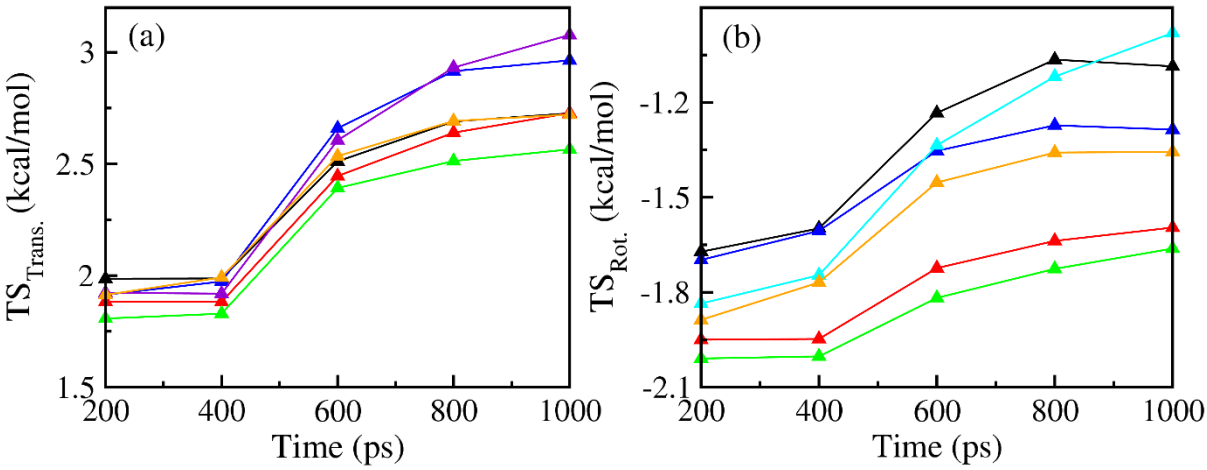

Figure 6.4 The convergence of (a) translational (TS_{Trans}) and (b) rotational ($TS_{Rot.}$) entropy values is shown by evaluating the entropy values at different time interval for pure water at 190 K. The different colors represent values for different water molecules. The lines are used to guide the eyes.

From Fig. 6.3, it can be seen that similar to the dynamic transition known to occur in water, both $TS_{Trans.}$ and $TS_{Rot.}$ values also show a transition at a particular temperature. The dynamic transition that takes place in water is known to be accompanied by a change in structure. Evidences for growing structures have been shown in water at a low temperature by showing increase in water's tetrahedral arrangement[45] or from the increase in four coordinated H-bonded water molecules[22]. These features make the dynamic motion taking place in water cooperative in nature, resulting in the hopping mediated dynamic processes.[22] From our study, we find that these processes also have their effect on the entropic behavior of water as the molecules show deviation from linearity in the entropy values at low temperatures. The results

show that even after the long simulations, the entropy values have a change in their pattern after the dynamic arrest. The likely explanation for this behavior could be that even in this timescale, the water molecules do not access all its possible configurations, resulting in a abrupt lowering of entropy values similar to the diffusion coefficient values.

Since we have both the individual water entropy values and diffusion coefficients at different temperatures, we wanted to investigate if a relation between the dynamic and thermodynamic properties of water exists. If such a relation exists, it will benefit us by calculating a computationally easier quantity (diffusion coefficient) and relate to a more demanding but important one (entropy). In general, the Adam-Gibbs (AG) relation of the following form is used to relate the configurational entropy (S_{Conf}) and the diffusion coefficient [46]:

$$D^{-1}(T) = D^{-1}(T_0)\exp\{A_{AG}/(TS_{Conf})\}, \qquad (6.3)$$

where $D(T_0)$ is the diffusion at a higher temperature T_0. A_{AG} is a constant known as Adam-Gibbs parameter. For our purpose, we used an analogue of Eq. 6.3 to fit $S_{Trans.}$ and diffusion where the term S_{Conf} has been replaced by $S_{Trans.}$ in Eq. 6.3. Fig. 6.5 shows the plot of the logarithm of inverse D with respect to $TS_{Trans.}$ values above the FSC transition temperature (240 K). From the plot, it can be seen that the diffusion coefficient and the entropy values are highly correlated (correlation coefficient of 0.98) indicating that there exists a general relation between the two quantities.

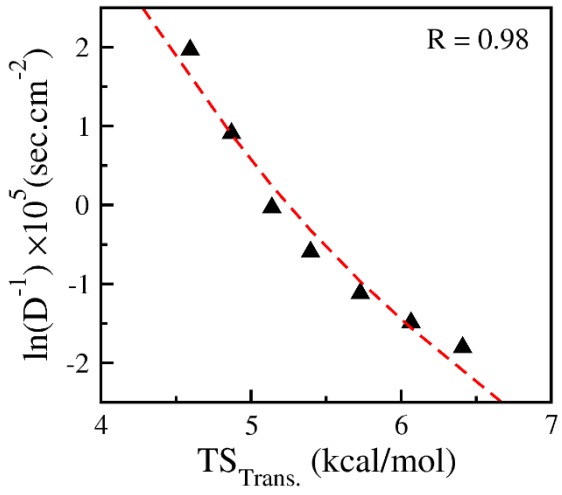

Figure 6.5 $\ln(D^{-1})$ vs. TS_{Trans} plot for individual water molecules. The red dashed lines here shows the fit with an analog of AG relation. The R is the Pearson's correlation coefficient.

After establishing the relation between the dynamics and entropy of water, we now investigate the entropic behaviour of individual water molecules in more detail. To understand the reason for the change in the entropy behaviour with temperature, , the individual water entropy values at various temperatures for both $TS_{Trans.}$ and $TS_{Rot.}$ are plotted in Fig. 6.6. From the Fig. 6.6(a), we observe that at higher temperatures the water molecules in the same system have similar $TS_{Trans.}$ Values. However, below the FSC transition temperature, different water molecules show different entropy values although being in the same solution. Since $TS_{Trans.}$, in our calculation, corresponds to the accessible volume available for a water molecule at a particular position, different entropy values indicate that the water molecules at low temperature experience different accessible volumes from each other. This implies that water structure at low temperature is heterogeneous in nature. At higher temperatures, it is likely that the environment near different water molecules change rapidly resulting in similar values of $TS_{Trans.}$ for all water molecules. In case of $TS_{Rot.}$, the variation in values are very smaller above FSC transition temperature compared to $TS_{Trans.}$ values as can be seen from Fig. 6.6(b). This indicates that $TS_{Rot.}$ values have a lesser dependence on temperature compared to $TS_{Trans.}$ for water. However, a clear change appears for $TS_{Rot.}$ values at lower temperatures. Similar to $TS_{Trans.}$ values, the $TS_{Rot.}$ values also show variation for water in the same solution. These patterns show that at lower temperature, the properties of water molecules start to differ from water at higher temperatures at the individual molecule level. A possible cause for such behavior could be the structural features of water, which can alter at different conditions. Our evidences indicate the presence of structural heterogeneity for water at supercool state, the exact origin for the observed pattern has been discussed in the next section.

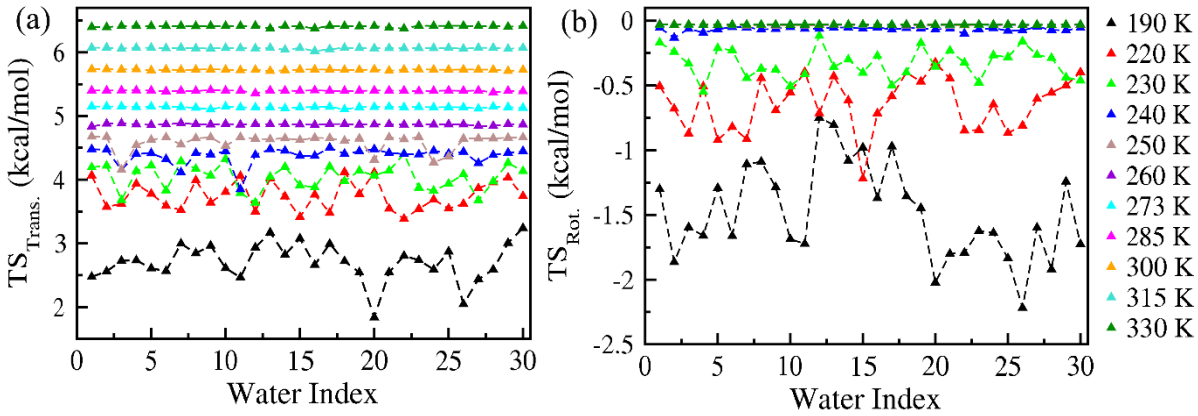

Figure 6.6 The individual $TS_{Trans.}$ (a) and $TS_{Rot.}$ values (b) for water at different temperatures. The dashed lines are used to guide the eyes.

6.3.2 Relation between Water Structure and Entropy

Water is known to exhibit structural polymorphism at supercooled states. There exists numerous structurally different forms of water with similar energies.[47] Therefore, it is likely that different structures of water may control the entropic behavior observed in our study. However, the distinction between the two most prominent water structures: the LDL and HDL water are very narrow. Hence, we need a parameter that can differentiate the microscopic properties characteristics of these water structures. For this purpose, we have used the structural order parameter, ζ reported by Russo and Tanaka[24] that distinguishes the HDL and LDL water. The parameter ζ is shown in Fig. 6.7(a) and is defined as the distance between the furthest water that makes H-bond with the 0^{th} water and the closest water not H-bonded to the 0^{th} water. For water molecules that are present in a high density environment, the value of ζ is expected to be small, while for water molecules that are more ordered with maximum number of H-bonds, the value of ζ would be higher. In the study by Russo and Tanaka[24], this parameter represents the average value of all water molecules, thus depicting the overall density of the solution. However, we can look at the properties of individual water molecules. Therefore, in our study, we look at the density environment of a particular water molecule present at a particular position near the ions, and we do so at different temperatures.

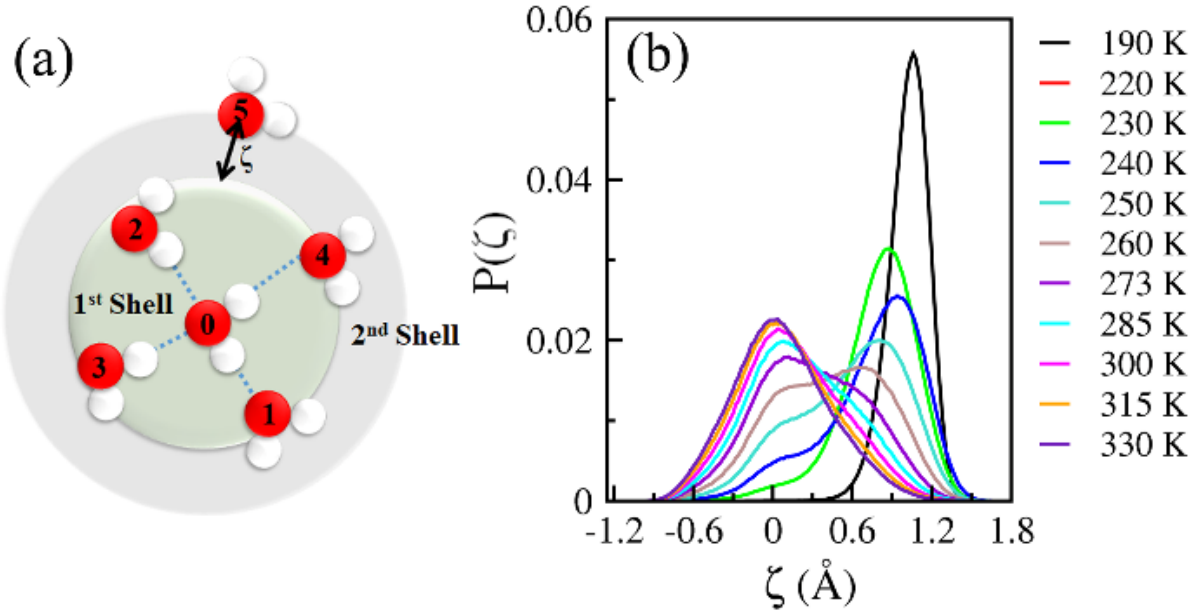

Figure 6.7 The schematic representation of the order parameter ζ used in this study (a). The water 1-4 are shown to form H-bonds with water "0" as shown by the blue dotted lines. The imaginary sphere within which these water molecules reside are shown to be 1^{st} (solvation) shell. The water molecule marked "5" represents the closest water that resides outside the 1^{st} shell and do not form H-bond with

water "0". The distance between the 1st shell and the water marked 5th in the 2nd shell represents the value of ζ. (b) The distribution of ζ for one of the representative bulk water molecule at different temperatures is shown.

In Fig. 6.7(b) the value of ζ at different temperatures has been shown for a randomly chosen water molecule from the bulk water system. We find here that the values of ζ has a very similar behavior as shown in Ref. 24 for the overall bulk water system. While at the lowest temperature studied, 190 K, the peak appears around 1 Å, for the highest temperature, 330 K, the peak comes close to 0 Å. This indicates that at higher temperature the water exists mostly as a HDL while at a very low temperature, it exists primarily as LDL. However, at an intermediate temperature, 260 K, a bimodal distribution appears peaking at both the two above values indicating a coexistence of both HDL and LDL. Interestingly, however, these structural coexistences do not affect the entropic behavior of individual water molecules as all water molecules at this temperature show similar $TS_{\mathrm{Trans.}}$ values (Fig. 6.6). Yet, at low temperatures, where we primarily observe LDL, water molecules show different entropic behaviour. Note that, we have ruled out the possibilities of non-convergence of entropy values by simulating the system up to a microsecond for the systems at lower temperatures. Therefore, assuming no further changes possible, we find an interesting observation that at the lower temperature, bulk water shows different possible values of entropy.

Therefore, to see if structural polymorphism affect the water entropic behaviour, we further investigated the relation between the structural parameter ζ and the individual entropy values at different temperatures. First we categorized water molecules into two subsets: high entropy water (HEW) and low entropy water (LEW) based on whether they fall above or below the average. Then, the average ζ values for water molecules having higher and lower entropy values were calculated for a total of 100 water molecules and the distribution of ζ, P(ζ), was plotted for these two categories at two different temperatures in Fig. 6.8. The distribution of ζ at 190K is shown in Fig. 6.8(a). From the figure, we see that water molecules with different entropy values have different distribution of ζ: water with slightly higher $TS_{\mathrm{Trans.}}$ values (red graph) remain in slightly higher density environment (lower ζ) while water with slightly lower entropy (black graph) remain at slightly lower density environment (higher ζ).

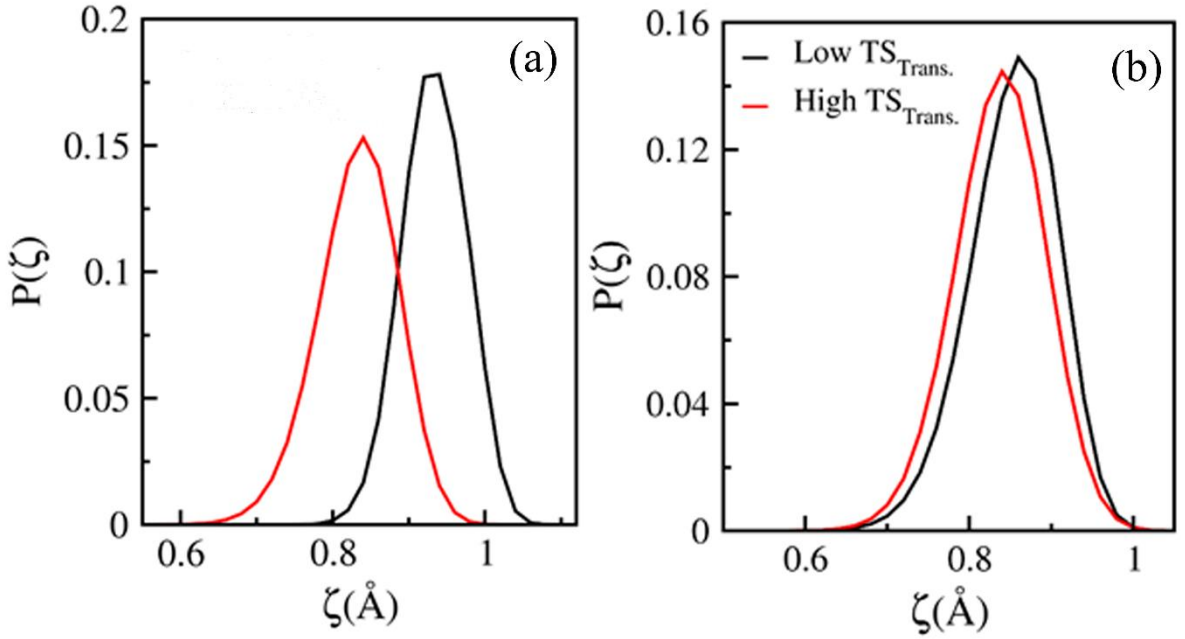

Figure 6.8 The distribution of ζ for water with higher (red line) and lower (black line) entropy water compared to the average $TS_{Trans.}$ for pure water at 190 K (a) and 220 K (b).

The pattern observed in Fig. 6.8(a) provides a clear evidence for structural heterogeneity of water molecules at low temperatures. This indicates that the different entropic behaviour is not just due to sampling of the configuration, but an inherent nature of the system itself at the lower temperature. Because of such structural heterogeneity, differently structured water molecules possess different entropy values. At 220 K, when the distribution of ζ is plotted for water with high and low entropy values in Fig. 6.8(b), the peaks do not show as clear distinction as observed at 190 K (Fig. 6.8(a)). This indicates that the spatial heterogeneity present in water become more prominent at lower temperatures only. However, the fact that at 220 K (even a little at 230 K also) entropy values are dissimilar to each other indicates that the spatial heterogeneity exists in case of water at these temperatures as well.

With the evidence of spatial heterogeneity present in water at low temperatures, the question arises is whether the HDL and LDL water co-exist in the same solution at these conditions. To address this, we have looked at the population of HDL and LDL water at different temperatures. However the definition of the two kinds of water involves some arbitrariness and may change the observed population. For our purpose, we have defined the HDL and LDL based on the value of ζ. To define HDL water, the average $\langle\zeta\rangle_{HT}$ and its standard deviation $\pm\delta\zeta_{HT}$ at 300 K is calculated. Hence, the water molecules possessing values of $\zeta < (\langle\zeta\rangle_{HT} + \delta\zeta_{HT})$ is called the HDL. To define the LDL, the average and standard deviation of ζ is calculated at 190K and

water molecules with $\zeta > (\langle\zeta\rangle_{LT} - \delta\zeta_{LT})$ is taken to be a LDL water. This way, we arrived at the criteria that a water molecule is in HDL or LDL state if it has ζ value lower or higher than 0.32 and 0.65 Å, respectively. The water molecules possessing ζ values in between these two limits are called intermediate between HDL and LDL water. Using these definitions, we measured the percentage for different types of water present in the systems at different temperatures. The plot is shown in Fig. 6.9.

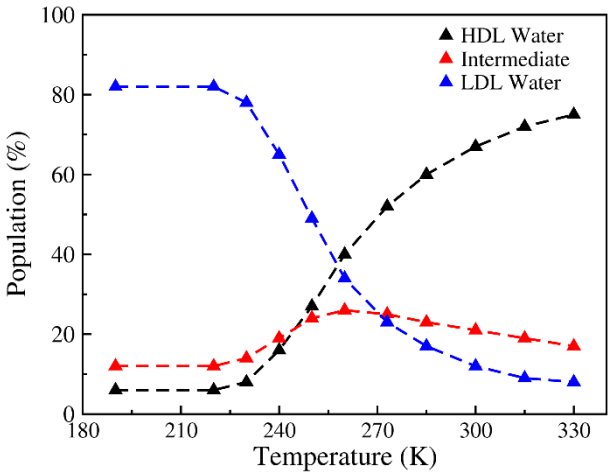

Figure 6.9 The population of HDL (black points) and LDL (blue points), and intermediate (red points) water at different temperatures. The dashed lines are used to guide the eyes.

From the figure, it can be clearly seen that at high or low temperatures, mostly one type of water molecules remain present predominantly. The two types of water have equal population only at temperature (~260 K) slightly above the FSC transition (240 K). Previously, Wikfeldt *et al.*[20] had shown the population of the different kinds of water at different temperatures. In their study, they find the co-existence for the two kinds of water at 230 K. However in their study, only the HDL and LDL water structures were considered using a particular value for the structure index without considering the intermediates. In our case, such strict definition is avoided to characterise the water molecules more appropriately. Also, the population for water belonging to neither HDL nor LDL state is low (<30%) at all temperatures. This also implies that it is either HDL or LDL water that are mostly present predominantly at all conditions. From Fig. 6.9, it is clear that although at low temperatures, we see existence of spatial heterogeneity in the water molecules, these water molecules do not correspond to the structural heterogeneity, i.e., HDL and LDL water. Rather, they belong to different kinds of water in the LDL state itself since both the peaks are above 0.65 indicating that both belong to LDL (Fig. 6.8(a)). Note here that the populations obtained in our study depends on the criteria used to characterise different water molecules. Hence, changing the criteria may alter the exact

population of these water molecules. The arbitrariness in assigning differently structured water has also been discussed by Moore and Molinero[21]. It has been stated that the evidence for structural polymorphism observed in different experimental studies also originate from such arbitrariness in assigning differently structured water molecules. From our analysis, we claim that the polymorphism reported in various experimental studies may originate from the spatial heterogeneity present in the LDL state of water only, which gives rise to various signals in experimental studies.[19-21] Hence the earlier belief that the different signals obtained from experimental studies correspond to co-existence of HDL and LDL water may not be correct. In the next section, we explore the origin for the spatial heterogeneity observed in supercooled water molecules.

6.3.3 Origin of Spatial Heterogeneity in Supercool Water

So far from our analysis we find that water molecules at low temperature possess different structural features giving rise to a spatial heterogeneity. To investigate how water molecules having different structural features are different from each other in terms of energy, we have carried out inherent structure analysis. 50 configurations have been collected from simulations at each temperatures and were energy minimized using conjugate gradient method to reach the closest local minima for these configurations. To see how these structures differ from each other, we first calculated ζ for individual water molecules from these configurations and plotted the average distribution for these ζ values. The distribution is shown in Fig. 6.10(a). From the figure, it can be seen that the inherent structures obtained from configurations at various temperatures have very small difference from each other in terms of ζ compared to the structures from higher temperatures. While a clear distinction for the peaks can be seen from the distribution in Fig. 6.7(b) at different temperatures, Fig. 6.10(a) shows that the configurations obtained from simulations at various temperatures possess a much smaller distinction. However, the difference in energy among these configurations are likely to originate from their small structural changes. Therefore, we evaluated the potential energy for the minimized structures obtained from different configurations. These values are plotted in Fig. 6.10(b).

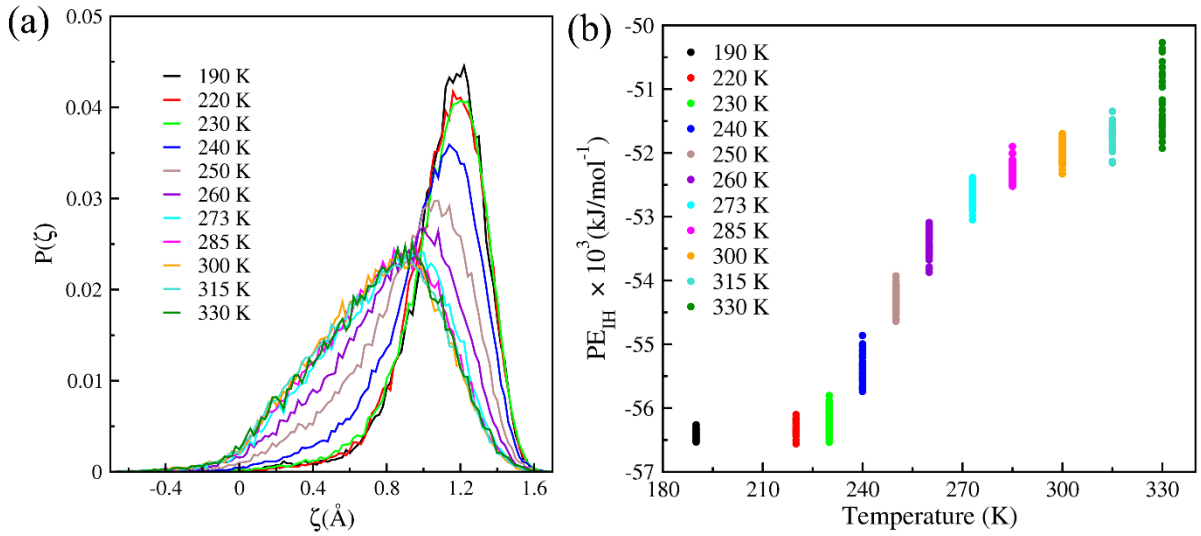

Figure 6.10 (a) The average distribution of ζ for water molecules from the energy minimized 50 configurations taken from simulations at different temperatures. (b) The energies of the inherent structures obtained from energy minimizing the configurations at different temperatures.

From the figure, it can be seen that configurations obtained from slightly higher temperatures has higher energy compared to configurations taken from low temperature simulations. For example, the configurations from 300 K and 190 K have nearly 5 kJ/mol difference in energy. Since the only difference between these configurations exist in terms of structure, the energy differences in them are likely to originate from such structural features only. The observed energy differences from Fig. 6.10(b) indicate that at low temperatures, water molecules remain stuck at different energy minima corresponding to particular value of ζ. The thermal energy available at these temperatures are not sufficient to overcome the energetic barrier corresponding to the structures with different ζ values.

To support the idea of existence of differently structured water with different energy values, we turn our focus to the barrier for going from one type of water structure to another. Hence, to measure the barrier for water to undergo its structural changes, the timescale for conversion between differently structured water is checked. To do this, we have measured the time taken for water to go from completely LDL state to completely HDL state. For this purpose, we consider a water molecule to be in LDL state if it has ζ value higher than the average ζ for water at 190 K while the water is considered to be in HDL state if the ζ value reaches lower than the average ζ value at 300 K. The values of ζ which correspond to neither of these correspond to the intermediate states of water. Note here that, in this case, the HDL and LDL water definitions are not same as the case with population calculation. The reason for this is in

this case, we aim to observe the conversion between two clearly different water structures. Although this change in criteria may affect the timescale obtained at different temperatures, our aim here is to obtain a qualitative, physical understanding of the conversion. Once again, by altering the criteria for LDL and HDL water, the observed timescale may change at different temperature. Hence the timescales reported here represent a qualitative feature for conversion between HDL and LDL water. Once we define the two states for our measurement of timescale, the following correlation function is calculated:

$$C(t) = 1 - < p_{LDL}(0).\, p_{HDL}(t) >, \qquad (6.4)$$

where p_{LDL} (0) and $p_{HDL}(t)$ represent the probability of finding a water molecule in HDL at time t and LDL state at $t = 0$, respectively. Thus the correlation function measures on average the time taken for a water molecule to convert from LDL to HDL state. Similarly, we can measure the reverse correlation where given a molecule in HDL state at t=0, it will convert to LDL at time t. The similar condition has been used in our previous study for water residence calculation around DNA base pairs[48-49]. The correlation function used here only focuses on water which are either in completely LDL or HDL state. The advantage of this correlation function is that a water molecule starting from LDL state has to reach the HDL state and therefore transient escape from LDL states are not considered in the measurement of timescales. The timescale has been obtained for conversion between the two structures is fitting the correlation function using an exponential fit. The plots could be fitted using a single exponential function for simulations above 250 K, while the curves below this temperatures require bi-exponential fits. The timescale obtained from the fitting are plotted with respect to temperature in Fig. 6.11.

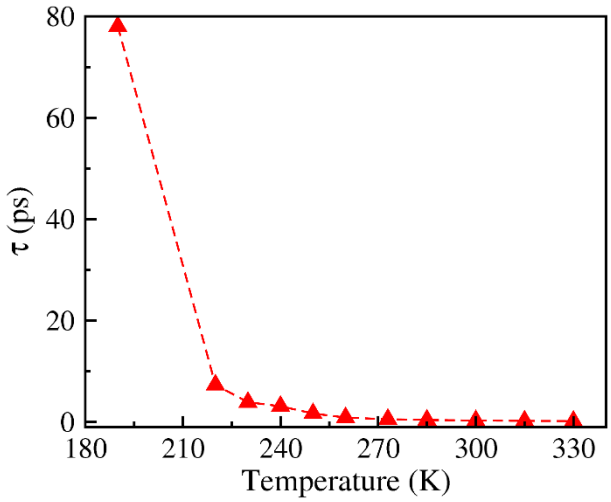

Figure 6.11 The time required for water molecules to convert from LDL state to HDL state at different temperatures. The dashed line is used to guide the eyes.

From the figure, it is clear that the conversion between differently structured water molecules are very fast at higher temperatures as expected. However, this timescale starts to increase as the temperature is lowered. An interesting observation that emerges from the figure is that the time taken for a water molecule to reach HDL state from LDL state is significantly higher at 190 K compared to the timescales observed at any other temperatures. This indicates that the barrier present for conversion between differently structured water at 190 K is significantly higher compared to water at any other temperatures. This also explains why we see clearly distinct peaks at 190 K for water with different entropy values since the changes in structure at this temperature is more difficult compared to other temperatures studied here. However, the distinction between the distributions of ζ for water with different entropy values are not very clear at higher temperatures compared to 190 K. This suggests that even at 220 K, water molecules can easily convert between HDL and LDL state using its thermal energy. At lower temperatures, the timescales and consequently the barriers for conversion between differently structured water increases significantly. These evidences clearly prove that water molecules at these low temperatures remain in a rigid, spatially heterogeneous environment. Hence, while talking about the properties of water molecules at these conditions, a more accurate approach should be to focus on properties of individual water molecules, rather than the whole system.

6.4 Conclusions

In this part of the thesis, the supercooled state of water has been studied through measurement of entropy of individual water molecules. From the calculation of translational and rotational entropy values of water at various temperatures and atmospheric pressure, we find new details on the behavior of supercooled water that remained unknown in the previous studies in this field. The calculated entropy values of individual water molecules were found to undergo a transition at similar temperature where they show a transition in their dynamic properties. Also, below the FSC transition temperature, individual entropy values from the same system were found to show diverse values. An investigation of local structures of water reveals that water molecules possess spatial heterogeneity in its structure at low temperatures. This structural heterogeneity in water gives rise to different entropy values for individual water molecules. The origin of such structural heterogeneity has been explained through inherent structure analysis and by looking at the timescale for conversion between LDL and HDL water. Our study finds that differently structured water possess different energies. Therefore water molecules at low temperatures remain stuck at these local minima having different energy values and this results in different entropy values for individual molecules present in the same solution. Our study also finds that although differently structured water molecules co-exist at low temperatures, all these water molecules belong to the LDL state of water. From our study, we claim that the evidences obtained for co-existence of HDL and LDL water through different experimental studies originate from the structural variations present in supercool state. Our study also highlights the importance of probing the individual water properties at low temperatures to present a detailed microscopic picture of the supercool water. However, the present study naturally brings in a new question about how the properties of water discussed above for bulk water will be modulated in the solvation shell of different solutes at the supercooled state. This will be the topic for the final chapter of the thesis.

6.5 References

1. Angell, C. A.; Shuppert, J.; Tucker, J. C. *J. Phys. Chem.* **1973,** *77* (26), 3092-3099.
2. Speedy, R. J.; Angell, C. A. *J. Chem. Phys.* **1976,** *65* (3), 851-858.
3. Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press: New York, 1971.
4. Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*. Oxford Univ. Press: New York, 1969.

5. Angell, C. A. *Science* **1995,** *267* (5206), 1924.
6. Mishima, O.; Stanley, H. E. *Nature* **1998,** *396*, 329.
7. Pablo, G. D. *J. Phys. Condens. Matter* **2003,** *15* (45), R1669.
8. Speedy, R. J. *J. Phys. Chem.* **1982,** *86* (6), 982-991.
9. Poole, P. H.; Sciortino, F.; Essmann, U.; Stanley, H. E. *Nature* **1992,** *360*, 324.
10. Sastry, S.; Debenedetti, P. G.; Sciortino, F.; Stanley, H. E. *Phys. Rev. E* **1996,** *53* (6), 6144-6154.
11. Rebelo, L. P. N.; Debenedetti, P. G.; Sastry, S. *J. Chem. Phys.* **1998,** *109* (2), 626-633.
12. Pallares, G.; El Mekki Azouzi, M.; González, M. A.; Aragones, J. L.; Abascal, J. L. F.; Valeriani, C.; Caupin, F. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111* (22), 7936-7941.
13. Angell, C. A. *Science* **2008,** *319* (5863), 582.
14. Soper, A. K.; Ricci, M. A. *Phys. Rev. Lett.* **2000,** *84* (13), 2881-2884.
15. Pettersson, L. G. M.; Nilsson, A. *J. Non-Cryst. Solids* **2015,** *407*, 399-417.
16. Nilsson, A.; Pettersson, L. G. M. *Chem. Phys.* **2011,** *389* (1), 1-34.
17. Loerting, T.; Winkel, K.; Seidl, M.; Bauer, M.; Mitterdorfer, C.; Handle, P. H.; Salzmann, C. G.; Mayer, E.; Finney, J. L.; Bowron, D. T. *Phys. Chem. Chem. Phys.* **2011,** *13* (19), 8783-8794.
18. Gallo, P.; Amann-Winkel, K.; Angell, C. A.; Anisimov, M. A.; Caupin, F.; Chakravarty, C.; Lascaris, E.; Loerting, T.; Panagiotopoulos, A. Z.; Russo, J.; Sellberg, J. A.; Stanley, H. E.; Tanaka, H.; Vega, C.; Xu, L.; Pettersson, L. G. M. *Chem. Rev.* **2016,** *116* (13), 7463-7500.
19. Huang, C.; Wikfeldt, K. T.; Tokushima, T.; Nordlund, D.; Harada, Y.; Bergmann, U.; Niebuhr, M.; Weiss, T. M.; Horikawa, Y.; Leetmaa, M.; Ljungberg, M. P.; Takahashi, O.; Lenz, A.; Ojamäe, L.; Lyubartsev, A. P.; Shin, S.; Pettersson, L. G. M.; Nilsson, A. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (36), 15214.
20. Wikfeldt, K. T.; Nilsson, A.; Pettersson, L. G. M. *Phys. Chem. Chem. Phys.* **2011,** *13* (44), 19918-19924.
21. Taschin, A.; Bartolini, P.; Eramo, R.; Righini, R.; Torre, R. *Nat. Commun.* **2013,** *4*, 2401.
22. Moore, E. B.; Molinero, V. *J. Chem. Phys.* **2009,** *130* (24), 244505.
23. Clark, G. N. I.; Hura, G. L.; Teixeira, J.; Soper, A. K.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (32), 14003.
24. Russo, J.; Tanaka, H. *Nat. Commun.* **2014,** *5*, 3556.
25. Mallamace, F.; Branca, C.; Corsaro, C.; Leone, N.; Spooren, J.; Chen, S.-H.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (52), 22457.
26. Tracht, U.; Wilhelm, M.; Heuer, A.; Feng, H.; Schmidt-Rohr, K.; Spiess, H. W. *Phys. Rev. Lett.* **1998,** *81* (13), 2727-2730.
27. Pan, A. C.; Garrahan, J. P.; Chandler, D. *Phys. Rev. E* **2005,** *72* (4), 041106.
28. Stevenson, J. D.; Schmalian, J.; Wolynes, P. G. *Nat. Phys.* **2006,** *2*, 268.
29. Taborek, P.; Kleiman, R. N.; Bishop, D. J. *Phys. Rev. B* **1986,** *34* (3), 1835-1840.
30. Ito, K.; Moynihan, C. T.; Angell, C. A. *Nature* **1999,** *398*, 492.
31. De Marzio, M.; Camisasca, G.; Rovere, M.; Gallo, P. *J. Chem. Phys.* **2016,** *144* (7), 074503.

32. De Marzio, M.; Camisasca, G.; Rovere, M.; Gallo, P. *J. Chem. Phys.* **2017,** *146* (8), 084502.
33. De Marzio, M.; Camisasca, G.; Conde, M. M.; Rovere, M.; Gallo, P. *J. Chem. Phys.* **2017,** *146* (8), 084505.
34. Sasikala, W. D.; Mukherjee, A. *J. Phys. Chem. B* **2014,** *118* (36), 10553-10564.
35. Rick, S. W. *J. Chem. Phys.* **2004,** *120* (13), 6085-6093.
36. Schlick, T. *Molecular modelling and simulation: An interdisplinary guide* Springer: New York, 2010.
37. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.
38. Nosé, S. *Mol. Phys.* **1984,** *52* (2), 255-268.
39. Hoover, W. G. *Phys. Rev. A* **1985,** *31* (3), 1695-1697.
40. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.
41. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
42. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008,** *4* (3), 435-447.
43. Sastry, S. *Nature* **2001,** *409* (6817), 164-167.
44. Chen, S.-H.; Mallamace, F.; Mou, C.-Y.; Broccio, M.; Corsaro, C.; Faraone, A.; Liu, L. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (35), 12974-12978.
45. Kumar, P.; Buldyrev, S. V.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (52), 22130.
46. Adam, G.; Gibbs, J. H. *J. Chem. Phys.* **1965,** *43* (1), 139-146.
47. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. Cambridge University Press: Cambridge, 2013.
48. Saha, D.; Supekar, S.; Mukherjee, A. *J. Phys. Chem. B* **2015,** *119* (34), 11371-11381.
49. Saha, D.; Kulkarni, M.; Mukherjee, A. *Phys. Chem. Chem. Phys.* **2016,** *18* (47), 32107-32115.

Chapter 7

Supercool Water in the Solvation Shells: Similarity and Dissimilarity with Bulk Water

7.1 Introduction

The structural polymorphism of water at different conditions has been a matter of debate. In the previous chapter, our results showed that either the high density liquid (HDL) or the low density liquid (LDL) water, that remain in abundance at a particular temperature except the temperature close near fragile to strong crossover (FSC) transition. Also, we have found that there exists dissimilar structure in the LDL region itself. However, an important factor that adds to the complexity of water properties is that naturally occurring water is always found as a mixture of one or more solutes. Therefore, an absolute clear picture of water properties only arises when both neat and solvation shell water behavior are completely understood. In this regard, aqueous solutions of ions have been subjected to numerous studies. These systems not only present the simplest realistic scenario possible, the experimental studies of these ionic solutions are often advantageous since water can be cooled to lower temperatures without crystallizing to ice,[1-2] the origin for which has been discussed previously.[3-4] Also, the aqueous solutions of ions manifest several properties that are very similar to neat water. The evidence for structural polymorphism in aqueous solutions have been verified by different simulation studies[5-6] and experiments.[7] The Raman spectroscopy measurements performed by Suzuki and Mishima[8] on glassy solutions of salts like LiCl, NaCl and KCl showed evidence for two distinct OH stretch modes, indicative of two different water structures. The effect of concentration variation on the structural properties of water has shown the shrinkage of LDL water region with increasing the salt concentration.[9] Further, the existence of liquid-liquid critical point (LLCP) has been shown in the mixtures of water-glycerol[10-11] and also in NaCl water solution.[12] The dynamics of aqueous solution also have been shown to exhibit properties of bulk water with the evidence of fragile-to-strong crossover transition found for NaCl solution.[13] In view of these observations, the extent of microscopic similarity and dissimilarity between bulk water and solvation shell water remains a matter of debate.

In this part of the thesis, we focus on the behavior of solvation shell water around ions. We intend to probe the consequences of water's structural and dynamic features at different conditions by looking at the translational ($TS_{Trans.}$) and rotational entropy ($TS_{Rot.}$) of

individual water molecules in the vicinity of ions. The single water entropy calculation method has been employed here also for this purpose. The objective of our study is twofold. First, we investigate how the FSC transition affects the entropic behavior of water in the vicinity of ions when temperature is varied at fixed atmospheric pressure. Second, we explore to what extent the presence of ions dictate water behavior at the supercooled state. The patterns observed from our calculations have been analysed by looking at the local density environment of individual water molecules. Our study finds new information on the proposed structural polymorphism of water and its effect on water entropy, similar to the neat state of water in the vicinity of ions. The results also would help us to explain changes in water freezing pattern in presence of ions. Our study thus aims to enhance our understanding of solvation shell water at the supercooled state at a microscopic level.

7.2 Computation Details

All atom molecular dynamics simulations were performed in our study to find the properties of water at different conditions. We have taken systems with 1:2000 ion-water ratio in cubic simulation box. For the solvation shell of different ions, we have taken one NaCl and one CsI in 2000 water molecules. To prevent the mixing of cation and anion solvation shells, we have kept a distance restraint between the cation and anion throughout the simulations. We have used TIP5P/E water model[14] for its accurate density patterns at different temperatures. For the ions, we have taken recently developed parameters by Satarifard *et al.*[15] that are compatible with TIP5P water model. All the systems were first energy minimized using steepest descent method[16]. This was followed by heating the systems up to 300 K using Berendsen thermostat[17] using a coupling constant of 0.2 ps. The systems were then equilibrated for 2 ns at 300 K and 1 bar pressure using Nosé-Hoover thermostat[18-19] and Parrinello-Rahman barostat[20] with a coupling constant of 0.4 ps for both. Similar to Chapter 6, here also the equilibrated systems thus obtained were then subjected to heating or cooling up to the temperatures at which the simulations were needed to be carried out. In this study, we have performed simulations at 190, 220, 230, 240, 250, 260, 273, 285, 300, 315 and 330 K. Therefore, the systems were equilibrated at these temperatures using the above mentioned thermostat and barostat. The equilibration for systems below and at 273 K was carried out for 30ns while for systems above this temperature were subjected to 10ns of equilibrations. A final simulation under NPT condition was carried out for each system with a time step of 2 fs and frames saved at every

0.1 ps from which the entropy values and other properties were calculated. These trajectories were 1000 ns long for systems at 190, 220 and 230 K while for systems simulated between 240 K and 273 K, the trajectories were 200 ns long. The systems above these temperatures were simulated for 100 ns within which converged entropy values were obtained. The electrostatics were calculated with PME electrostatics[21] with a cut-off length of 10 Å for both electrostatics and van der Walls interactions. All the simulations were performed using GROMACS software package.[22] The method for entropy calculation has been discussed in Section 2.2.3 in Chapter 2.

7.3 Results

We first discuss how the variation of temperature changes in the single water entropy values in different solvation shells. The various aspects of water entropy and structure in the bulk and in solvation shell are discussed in this regard to show how ions affect water behavior at various temperatures. This will be followed by discussion on the attributes that make ionic solutions different from the bulk water.

7.3.1 Water entropy in different solvation shells

We begin the discussions by comparing the entropy values for water in the first and second solvation shells along with the values of bulk water entropy at different temperatures and normal pressure. Since our calculations involve evaluation of entropy values for individual water molecules, we first find out the number of water molecules in different solvation shells of ions by calculating the radial distribution function (RDF) of water oxygen around ions. Then we take the average of these individual entropy values for water that belong to the different solvation shells. Therefore, the average values here represent the entropy of a single water molecule present at a particular distance from the ion. The $TS_{Trans.}$ and $TS_{Rot.}$ values of water molecules are shown in Fig. 7.1 for first and second solvation shell water around Na^+, Cs^+, Cl^- and I^- along with the average values of bulk water entropy. As can be seen in Fig. 7.1, both $TS_{Trans.}$ and $TS_{Rot.}$ values show a deviation from linearity as the temperature is lowered for both bulk water and for water in the solvation shells. We have previously shown for neat water in Chapter 6 that the deviation in $TS_{Trans.}$ values took place at a similar temperature where the diffusion coefficient of water showed divergence. Interestingly, we find that the deviation for

bulk water and that for water in different solvation shells occur at nearly the same temperature. The deviation is clearer in case of $TS_{Rot.}$ values which occur below 240 K for all cases.

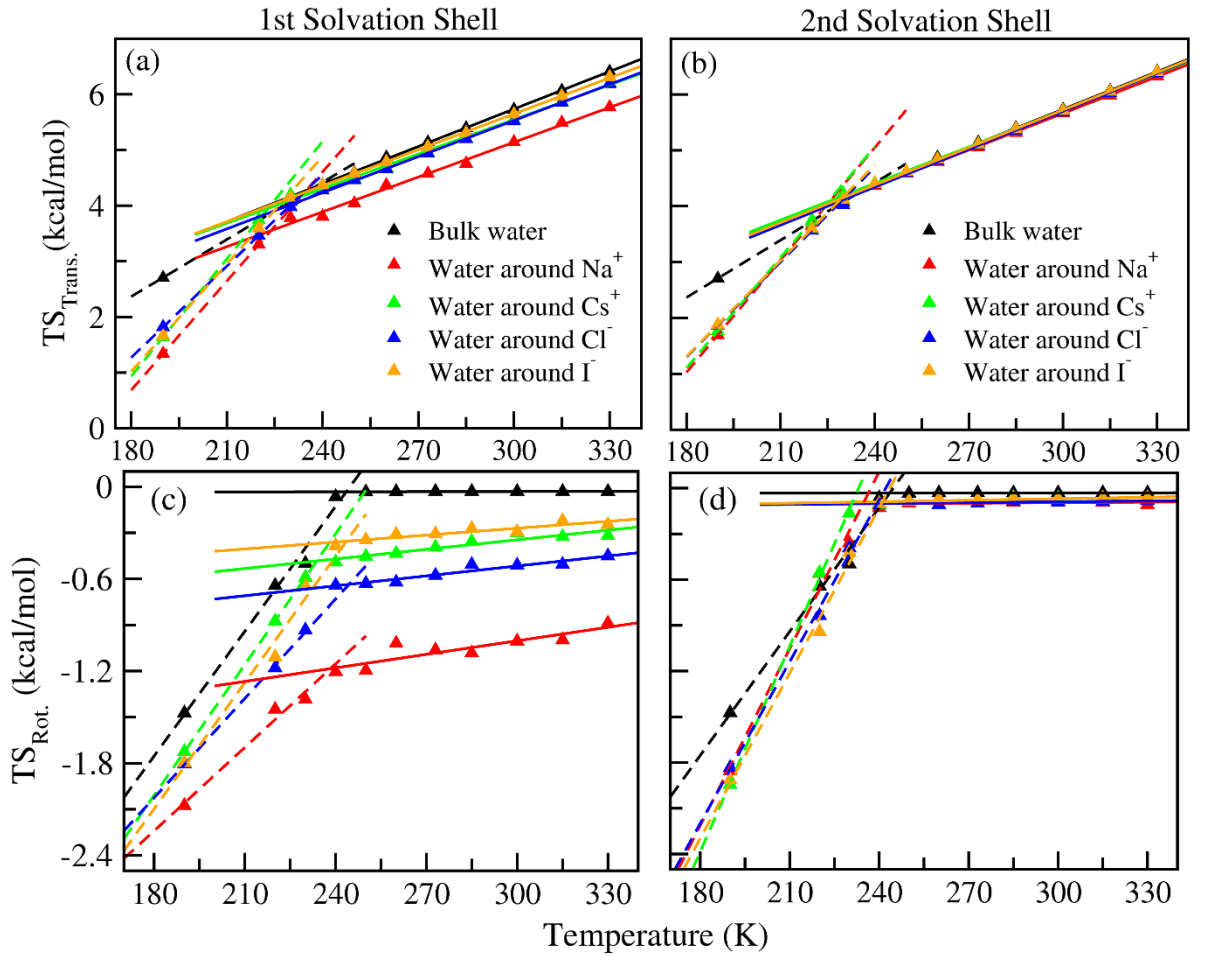

Figure 7.1 The single water translational, $TS_{Trans.}$ (upper panel) and rotational entropy, $TS_{Rot.}$ (lower panel) values for pure water and water in the solvation shells. The plots on the left show the values for first solvation shell and the plots on the right show values for second solvation shell of different ions. The lines here show the straight line fits of entropy values above (solid line) and below (dashed line) the FSC transition.

Gallo *et al.*[13] showed previously that the phenomena observed in bulk water are also seen in case of ionic solutions. In their study, the dynamic properties were measured at NaCl concentration of 0.67 mol/kg. However, to our knowledge, the properties of water in the solvation shell of ions have not been reported previously. Using our approach of single water entropy calculation, we find that the FSC transition have the same effect on water in both first and second solvation shell as that of the bulk water. The observation of this behavior for highly charged ion such Na^+ and low charge density ion like I^- indicate that the effect of water's FSC

transition transfers to the solvation shells in general. Next we shall discuss how the individual water molecules are affected in the bulk and in ionic solutions at supercooled states.

7.3.2 Individual Water Entropy Values

Since we find that the average entropy of solvation shell water shows similar temperature dependence as that of the bulk water, it is interesting to see how individual water entropy values vary at different temperatures. For this, we have plotted the $TS_{Trans.}$ values of individual water molecules near different ions at different temperatures. The values are shown in Fig. 7.2.

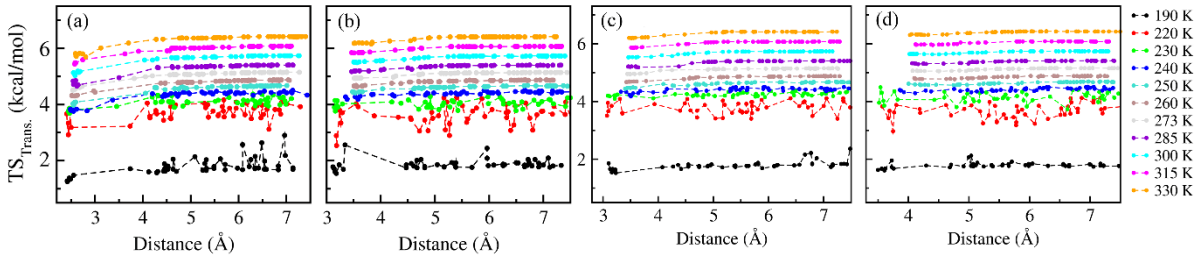

Figure 7.2 The individual $TS_{Trans.}$ values for water around (a) Na^+, water around (b) Cl^-, water around (c) Cs^+ and water around (d) I^- at various temperatures. The distances represent the average distance between the ion and the water oxygen atom. The lines here are used to guide the eyes.

From Fig. 7.2, we find that different molecules exhibit similar entropy values up to 240 K in different solvation shells. However, below this temperature, the individual entropy values start to vary from each other similar to the bulk water. Note here that the two points next to each other in Fig. 7.2 representing two water molecules do not correspond to water molecules that reside in the vicinity of each other in the simulation box. As seen in the case of average entropy of solvation shell water molecules, the individual water molecules also exhibit very similar pattern as that of bulk water for individual entropy values. Except for the few water molecules in the first solvation shell of ions, which are likely to be strongly affected by the electrostatic attraction of ions[23], we find that at lower temperatures, the water molecules display diverse values similar to our observation for bulk water. The same is observed in the case of water molecules around Cs^+ and I^- (Fig. 7.2(c) and 7.2(d) respectively) indicating the fact that these properties represent a general feature of water at supercooled state irrespective of its surroundings. Our finding also emphasises the fact that at these conditions, the thermodynamic behavior of different water molecules present in the same solution are distinct from each other. Hence, while investigating processes that take place at these conditions, consideration of the

characteristics of individual water molecules could prove more advantageous. We now look into the structural features of water that could be used to explain the observed water behavior.

7.3.3 Relation between Water Structure and Entropy

The structural features of pure water has been discussed in Section 6.3.2 in Chapter 6. The parameter ζ has been introduced to characterise the high and low density liquid (HDL and LDL) water. For pure water, we have shown that at different temperatures, only one kind of water exist in majority with the coexistence of two types of water only slightly above the FSC transition temperature. Since the individual water entropy values in the neighbourhood of ions show a similar pattern as pure water, it is interesting to investigate the structural features for water near different ions. Therefore, in order to check how the value for ζ varies for water in the solvation shell with high and low $TS_{Trans.}$ values, ζ has been calculated for individual water molecules in the bulk phase and in the vicinity of ions. Since the water in the first solvation shell of ions are likely to always have low value of $TS_{Trans.}$, ζ has been calculated from the second and higher solvation shells. To see the pattern of ζ for water with high and low $TS_{Trans.}$ values, the distribution of ζ for water molecules having higher and lower $TS_{Trans.}$ values with respect to the average $TS_{Trans.}$ of all the water molecules has been plotted for pure water and water near ions. Fig. 7.3 shows the distribution of average ζ for water with high and low $TS_{Trans.}$ values for water near Na^+, Cl^-, Cs^+ and I^-. For the ions, 40 closest water molecules have been considered beyond which the water behavior is indistinguishable from bulk water.

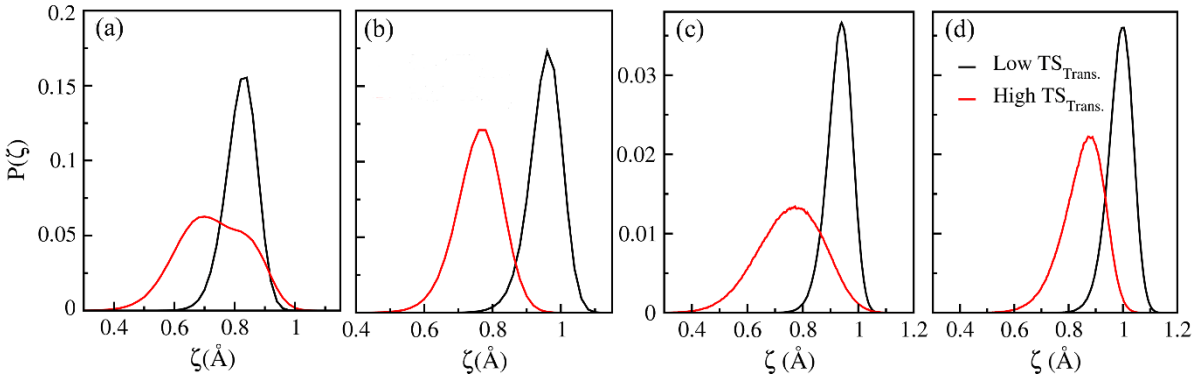

Figure 7.3 The distribution of ζ at 190 K for water near Na^+ (a), water near Cl^- (b), water near Cs^+ (c) and water near I^- (d). The red and black lines represents ζ for all the waters having lower and higher value of $TS_{Trans.}$ respectively compared to the average $TS_{Trans.}$ of all water molecules considered.

Form Fig. 7.3, a similar pattern as the bulk water emerges for all cases. The water molecules having different kinds of $TS_{Trans.}$ values have two different peaks as seen from the red and

black curves. The water molecules with higher $TS_{Trans.}$ values remain at a slightly higher density environment while the local density for water with low $TS_{Trans.}$ are found to be lower. Interestingly, this behavior is consistent for water near all four ions considered in our study as seen from Fig. 7.3(a)-(d). For systems at 220 K, this clear distinction is not found, indicating that these structural features and their consequence on individual entropy values become prominent at very low temperatures only. However, similar to the pure water, the diverse values of $TS_{Trans.}$ seen in the systems simulated at 220 and 230 K also indicate that the small structural variations could be responsible for such behavior at these temperatures as well. Once again from the positions of the peaks in Fig. 7.3, it can be seen that water molecules in these systems remain in LDL state and the variations in their structural properties do not correspond to coexistence of HDL and LDL water.

7.3.4 Role of Ions in Water Freezing

Presence of salt is known to decelerate ice nucleation by altering ice-liquid interfacial free energy.[4] However, we find that water even in the vicinity of ions show neat water like FSC transition in entropic behavior and very similar structural properties. Therefore, we turn our attention to the structural features of water very close to the ions. To see how the density of water molecules present in the first solvation shell of ions changes with temperature, the distribution of ζ has been checked for the closest water from different ions. The plots for the water near Na^+ and Cl^- are shown in Fig. 7.4. The figure shows that unlike the case of pure water, these molecules do not have distinct peaks at higher and lower temperatures. Although the peak position changes slightly at lower temperatures, the value never goes to such an extent where LDL water peaks are seen for the case of pure water at 190 K. This pattern, visible for both Na^+ and Cl^- indicates that the water molecules very close to these high charge density ions have a different behavior in terms of local density at various temperatures compared to pure water. Also, the structure of the water molecules near these ions are very similar to the HDL water seen for the case of pure water at higher temperatures. Thus, we wanted to probe whether such HDL structure of water are always present in the first solvation shell of different ions or not. To answer this, the distributions of ζ is plotted for the closest water near Cs^+ and I^- also and are shown in Fig. 7.5. Interestingly, we find that even the closest water from the ion also in this case shows very similar bulk like pattern at different temperatures. To better represent these patters, the mode for the distribution of ζ has been plotted for the closest water from ions

and for bulk water at different temperature. This is shown in Fig. 7.6. The plot clearly presents the changes observed for the water near Na^+ and Cl^- compared to bulk water or for water near Cs^+ and I^-.

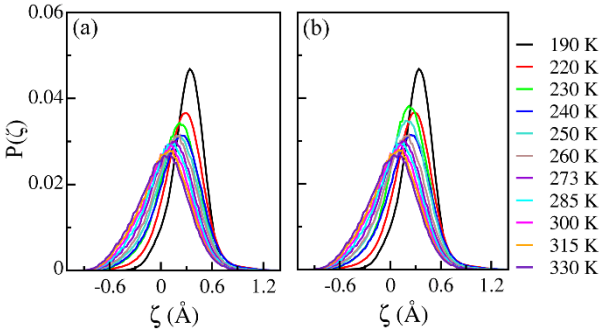

Figure 7.4 Distribution of ζ for the closest water from Na^+ (a) and from Cl^- (b) at different temperatures.

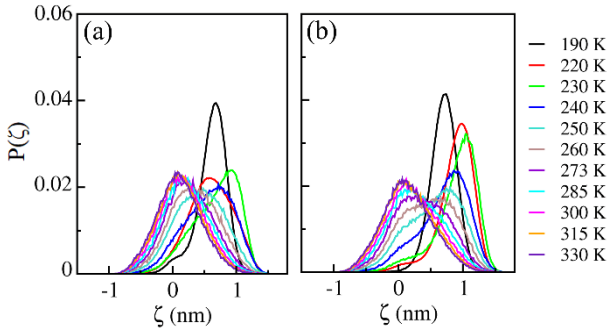

Figure 7.5 Distribution of ζ at different temperatures for the closest water from Cs^+ (a) and for the closest water from I^- (b).

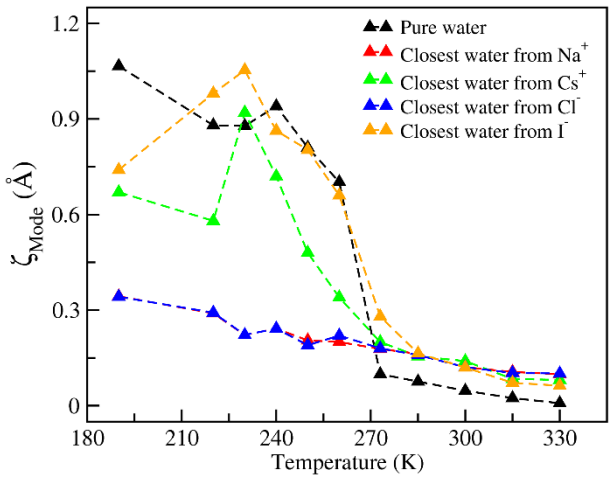

Figure 7.6 The value of ζ at the mode of ζ distribution plots for bulk water and for the closest water from Na^+, Cs^+, Cl^- and I^- at different temperatures. The lines have been used to guide the eyes.

Next, we checked how far the structural variations observed for the closest water from ions exist for other water molecules near different ions compared to the bulk water. To do this, we

have calculated the distribution of ζ for the 15th and 45th water with respect to distance from the all the ions. The plots are shown in Fig. 7.7 and 7.8, respectively. The figure shows a very similar behavior for all cases, and the pattern is found be very similar to the bulk water. Our findings indicate that the changes in structural features at different temperatures can only be seen for water residing in very close proximity of high charge density ions. Note here that the concentration considered in our study has been very low. Hence, with increase in the concentration of ions such as in case of sea water, it is likely that the solvation shells of different ions will mix with each other, resulting in more number of water molecules experiencing relatively temperature independent density environment.

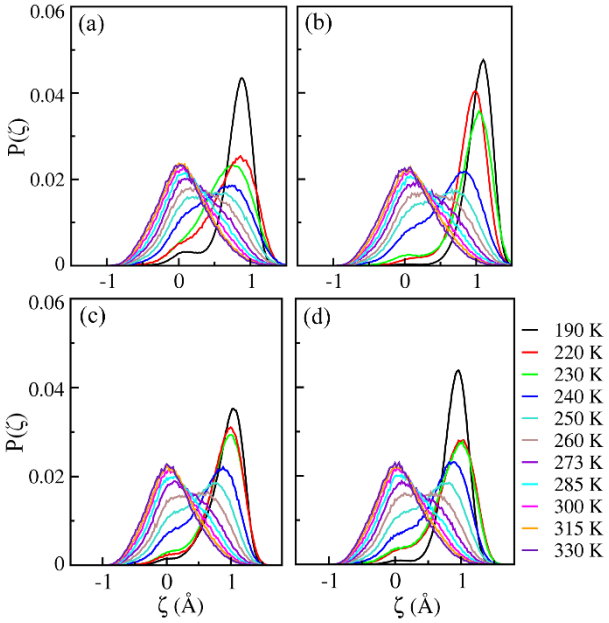

Figure 7.7 The distribution of ζ for 15th water molecule (upper panel) and 45th water molecule (lower panel) with respect to distance from Na^+ (a and c) and from Cl^- (b and d), respectively.

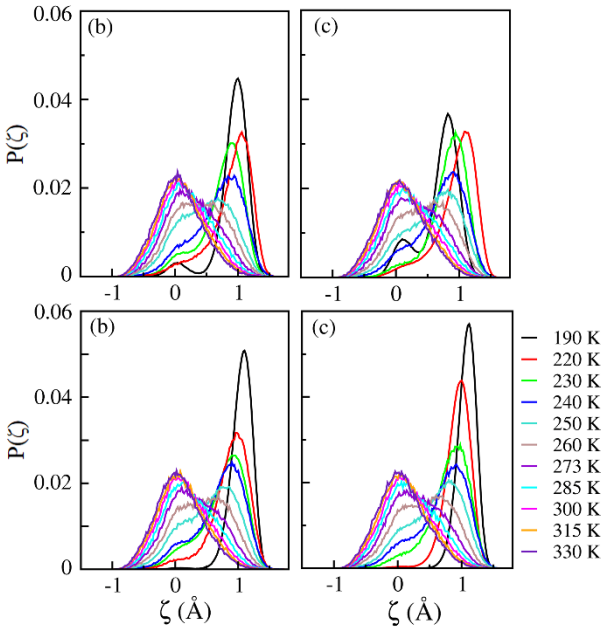

Figure 7.8 The distribution of ζ for 15th water molecule (upper panel) and 45th water molecule (lower panel) with respect to distance from Cs^+ (a and c) and from I^- (b and d), respectively.

To check how the water structure changes for water with increasing salt concentration, we have carried out two simulations with 1.36 M and 2.1 M concentration of NaCl as was done by Corradini *et al.*[9] to study NaCl solutions. We have taken 49 and 76 Na+ and Cl- ions for the two concentrations in 2000 water molecules. Sufficiently long equilibration has been carried out for different time periods at different temperatures and the average value of ζ for water molecules present in these solutions has been calculated from 10 ns long trajectories at each temperatures that has been considered in this study. The distribution of ζ for water at different temperatures are shown in Fig. 7.9(a) and 7.9(b) for the two concentrations. Interestingly we find that the distribution of ζ for water in these solutions are narrower compared to the bulk water. From this, we have looked at how ζ changes for water near Na^+ or Cl^- and in their solution compared to the bulk water by plotting the mode of ζ distribution in Fig. 7.9(c). We find that while the mode of ζ value for closest water molecules are far from bulk water values at lower temperatures, it shifts towards bulk water for average ζ values of water from the solutions.

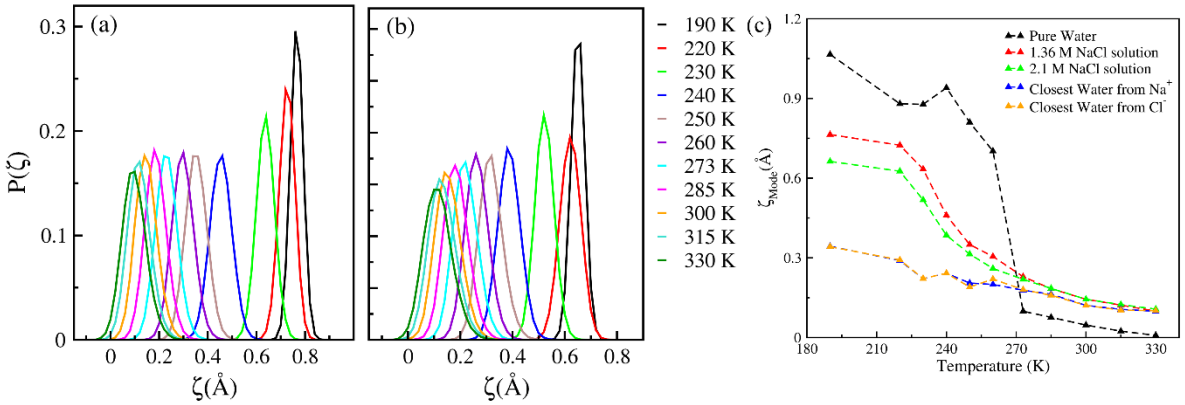

Figure 7.9 Distribution of ζ for all water from 1.36 M NaCl solution (a) and 2.1 M NaCl solution (b) at different temperatures. (c) The mode of ζ distribution for bulk water, water closest to Na^+ and Cl^- and from average ζ plot from different concentrations.

However, even at 1.36 M concentration, the mode of ζ values do not reach the bulk water values at low temperatures, thus hindering the formation of ice with low density. Thus, for these solutions, even lower temperatures would be required to reach the LDL state and eventually to form ice. Also, the fact that sea water has higher concentration of NaCl compared to other salts probably indicates why sea water does not freeze easily at places where water is likely to remain in ice form. Our observations are consistent with the findings of Corradini *et al.*[9] where the LDL region was found to shrink in presence of NaCl in the medium. However, here using our highly dilute solution and investigating the individual water behavior, we find out the exact extent of the ions' effects on water structure. Overall, our study shows that the ions' effect on water structural properties are very limited, with only high charge density ions being able to impart some changes in the behavior of a few of its closest water molecules.

7.4 Conclusions

In this part of the thesis, we have attempted to understand the influence of different ions on thermodynamic properties of water through measurement of entropy and structure. The translational and rotational entropy of individual water molecules have been calculated for water in the solvation shell of ions at various temperatures. We have obtained three main conclusions from our study. First, we find that the dynamical arrest known to take place at supercooled water gives rise to a transition in the entropy values also in different solvation shells. This behavior is true for the water molecules in the close vicinity of ions as well. This shows the dominance of bulk water properties for water at different environments at different

temperatures. This dominant behavior of water also explains why biomolecules show dynamic transition[24-25] at same temperatures where water undergoes its FSC transition. Second, our study shows that below FSC transition temperature, water molecules belonging to the same solution show diverse entropy values even in the solvation shells. Analysis of local structure showed that the origin of this behavior comes from co-existence of two different kind of water in the same solution. Our results show that the two types of water structure are not the HDL and LDL water, rather they represent variations of LDL water at low temperatures. Once again the patterns are consistent across all the ions considered and follow the pure water behavior.

Finally, we show that only a few water molecules residing very close the high charge density ions show some variations in their structural features from bulk water across various temperatures. For low charge density ions, even the closest water from the ion show bulk like behavior. We show that these small structural changes, which are likely to take place for more number of water molecules at higher concentration of ions is the reason why ions alter the freezing behavior of water. Overall, our study brings newer microscopic details to the understanding of water behavior in the solvation shells of ions at various conditions.

7.5 References

1. Miyata, K.; Kanno, H.; Niino, T.; Tomizawa, K. *Chem. Phys. Lett.* **2002,** *354* (1), 51-55.
2. Miyata, K.; Kanno, H. *J Mol Liq.* **2005,** *119* (1), 189-193.
3. Koop, T.; Luo, B.; Tsias, A.; Peter, T. *Nature* **2000,** *406*, 611.
4. Espinosa, J. R.; Soria, G. D.; Ramirez, J.; Valeriani, C.; Vega, C.; Sanz, E. *J. Phys. Chem. Lett.* **2017,** *8* (18), 4486-4491.
5. Paschek, D. *Phys. Rev. Lett.* **2005,** *94* (21), 217802.
6. Le, L.; Molinero, V. *J. Phys. Chem. A* **2011,** *115* (23), 5900-5907.
7. Ruiz, G. N.; Bove, L. E.; Corti, H. R.; Loerting, T. *Phys. Chem. Chem. Phys.* **2014,** *16* (34), 18553-18562.
8. Suzuki, Y.; Mishima, O. *Phys. Rev. Lett.* **2000,** *85* (6), 1322-1325.
9. Corradini, D.; Rovere, M.; Gallo, P. *J. Phys. Chem. B* **2011,** *115* (6), 1461-1468.
10. Suzuki, Y.; Mishima, O. *J. Chem. Phys.* **2014,** *141* (9), 094505.
11. Murata, K.-i.; Tanaka, H. *Nat. Mater.* **2012,** *11*, 436.
12. Corradini, D.; Gallo, P. *J. Phys. Chem. B* **2011,** *115* (48), 14161-14166.
13. Gallo, P.; Corradini, D.; Rovere, M. *J. Chem. Phys.* **2013,** *139* (20), 204503.
14. Rick, S. W. *J. Chem. Phys.* **2004,** *120* (13), 6085-6093.
15. Satarifard, V.; Kashefolgheta, S.; Vila Verde, A.; Grafmüller, A. *J. Chem. Theory Comput.* **2017,** *13* (5), 2112-2122.

16. Schlick, T. *Molecular modelling and simulation: An interdisplinary guide* Springer: New York, 2010.
17. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.
18. Nosé, S. *Mol. Phys.* **1984,** *52* (2), 255-268.
19. Hoover, W. G. *Phys. Rev. A* **1985,** *31* (3), 1695-1697.
20. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.
21. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
22. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008,** *4* (3), 435-447.
23. Saha, D.; Mukherjee, A. *J. Phys. Chem. B* **2016,** *120* (30), 7471-7479.
24. Réat, V.; Dunn, R.; Ferrand, M.; Finney, J. L.; Daniel, R. M.; Smith, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2000,** *97* (18), 9961.
25. Rasmussen, B. F.; Stock, A. M.; Ringe, D.; Petsko, G. A. *Nature* **1992,** *357*, 423.

Chapter 8

Summary and Future Directions

8.1 Summary

The investigation of water properties has been a subject of research for last few decades. With the advancements of new experimental and simulation methods, it is likely that water will continue to unravel new properties at various environments and conditions. The thesis titled "*A Molecular Perspective of Dynamics and Entropy of Solvation Shell Water"* discusses some of the aspects of water, which sheds new lights on its behavior at local level. A major part of the study focuses on behavior of single water molecules, which have led to new information that has remained overlooked in different studies.

The dynamics of water near DNA base pairs have been studied through measurement of mean residence time of water in the solvation shell. We have focused on the chemistry of the base pairs and the geometric features of DNA structure to identify the factor that governs the dynamics of water molecules. Our study finds that rather than the chemical nature of the base pairs or the structural factors, it is the position of the base pair in the 12-base pair DNA systems that plays the most important role in determining the residence time values. From analysis of our results and through a simple kinetic model, our study finds that the water molecules diffuse to different sites in DNA before escaping the solvation shell, resulting in higher residence time at the middle of the DNA.

The residence time study has been extended to DNA in hydrated ionic liquid systems in the context of DNA's long term storage purpose. In this case, we found that CG and AT base pairs show very different behavior for the residence time of ionic liquid cations. This behavior stems from the different affinity of DNA base pairs toward water in these systems. The CG base pairs tend to interact more with its surrounding water, resulting in a crowded environment around itself. This results in slowdown in the dynamics of ionic liquid cations near these bases. Such difference in the dynamics and environment near different base pairs for ionic liquid cations has been discussed in the context of DNA stability in hydrated ionic liquids.

For liquid systems, the dynamics and their thermodynamic properties are often interdependent on each other. While the dynamics of water is important mostly from a functional point of view of biomolecules, the thermodynamics is crucial for the stability and interactions of different molecules. In this thesis, the thermodynamics of water has been studied through measurement of water entropy at individual molecule level. Using a method recently developed within our group, we have studied the entropy of individual water molecules in the solvation shell of different cations and anions. Through comparison with previously known entropy values, we have established the accuracy of our method. Further, the study finds the exact extent of entropy changes in the solvation shell of ions. Individual water entropy calculations reveal that while cations primarily modulate the translational entropy of water, the anions through their hydrogen bonding ability affect the rotational motions of water, and consequently their rotational entropy values. Overall, the study highlights the intricate differences in the way positive and negatively charged solutes modify water behavior.

Different conditions can impact the structure and dynamic properties of water molecules. In this context, we have investigated the effect of dynamic arrest known to take place in the supercooled water on individual water entropy values. Our aim has been to connect the thermodynamic properties of water, i. e. water entropy with its dynamics. Therefore, through calculation of diffusion coefficient and single water entropy at different temperatures in the pure water system, we have shown how these two quantities are related to each other. Then we have investigated the effect of dynamic arrest and structural polymorphism of water at different temperatures. We found that similar to the dynamics of water, both translational and rotational entropy values show a transition in their value at the same temperature where water shows a dynamics arrest. Also, below this temperature, we found that different water molecules present in the same system possess distinct entropy values. The analysis of water structure revealed there exist structural heterogeneity in the water molecules at low temperatures, which results in such entropic behavior. The results also showed that instead of co-existence of high density and low density water molecules, the structural heterogeneity is present in the low density liquid water structure itself. The results highlighted the significance of observing single water behavior at supercool state, which has not been explored previously.

Finally, the supercool water properties has been investigated in the solvation shell of ions to get a more realistic picture of water behavior. Interestingly, we found that the solvation shell water shows very similar behavior to the bulk water. The transition in entropy values has been found to take place at similar temperature for both first and second solvation shell water. Also, the structural heterogeneity observed in case of bulk water have been found to exist in the vicinity of solvation shell water also. In spite of such similarities, we have attempted to provide an explanation for the changes in freezing behavior known to exist in presence of ions. Through analysis of single water properties, we identified the exact changes which prevent water molecules to go to ice state, thus resulting in changes in the freezing behavior.

Overall, the thesis provides new insights on water dynamics, thermodynamics and structural features at different environments and conditions through a molecular approach.

8.2 Future Directions:

8.2.1 Dynamic Properties of Water

While this thesis talks about the dynamics of water near DNA base pairs, numerous scenarios exists where the residence time of water can bring important information regarding biomolecular functions and interactions. The success of our approach in identifying the principal factor involved in governing residence time of water indicates that the method is applicable for study of water molecules near proteins, ion channels or cell membranes. In addition to these, the approach can be extended to a number of cases which can not only lead to significant progress for biomolecular simulations, it can also help in developing new methods for more application oriented approaches. Few of these examples are discussed in the following.

(A) Water Dynamics as a Tool for Testing Biomolecular Force Fields.

In general, the commonly used force fields are far from perfect for representing the exact behavior of biomolecules. While majority of the force fields are only accurate for study of folded structures of proteins, the structures of intrinsically disordered proteins (IDPs) often deviate from their real structures in terms of radius of gyration[1]. In recent years, a large number of studies have been carried out to rectify this issue[2-5]. While some of these modified parameters alter the interactions

between water and protein atoms[2], others have modified the interactions between the protein atoms to come up with perfect structures for the IDPs[3]. Although such approaches have been successful in matching the properties of IDPs between experiments and simulations, the accuracy of such modified interaction parameters for folded proteins remain an uncertainty. Recently, Robustelli *et al.*[4] have been successful in developing parameters which is accurate for both folded and unfolded protein systems.

Although these studies have marked significant improvements in biomolecuar simulations, the properties of water molecules near the biomolecules remain an important aspect for the application of new force field parameters. Therefore, the applicability of these new parameters can come through accurate investigation of water dynamic properties. While scaling of interaction between water and protein atoms results in better structural agreement between experiment and simulations, it is also necessary to accurately investigate how such modifications affect the dynamics and thermodynamics of water molecules in the vicinity of proteins and peptides. These scenarios also bring the possibility of modifying the parameters of water for better compatibility with the new force field parameters. The fact that the approach used in the thesis provides timescales for water dynamics which are very close to experimental values, future studies can be carried out to establish the accuracy and viability of new force field parameters developed in recent times. Also, modifications can be obtained in the existing water models for more accurate picture of water dynamics and thermodynamics near IDPs.

(B) Calculation of Drug Residence Time near Biomolecules.

The applicability of newly developed drug molecules depend on several factors. One such factor is the duration of its interaction with a particular binding site, i.e. its residence time. In general, residence time values for different molecules have been widely used as a characteristic to identify the most suitable drug for different systems.[6] Several studies have investigated the time taken for a drug molecule to leave a particular binding site with different methods.[7-10]

The success of the approach used in our study to calculate mean residence time indicates that such simple formalism can be applied for the case of residence time measurements of drug molecules as well. However, in these cases, multiple drug molecules will be required to be present in the

simulation box so that multiple binding-unbinding events can take place in the simulation. To achieve such multiple binding-unbinding process, different approaches such as scaled molecular dynamics[7] can be applied. Also, the role of water dynamics in the drug binding sites can also be investigated to identify the role of water in more detail, which can improve the drug binding kinetics at a particular site. Overall, we believe that the success of our approach for DNA-water case can be extended to such complex systems also, which will lead to significant progress in the field of drug development.

8.2.2 Application of Single Water Entropy

In this study, we have obtained new information on the behavior of water entropy in its pure state and near different ions. However, the ions represent the simplest real system possible. In general, the biomolecules present a very different scenario from the ions due to the various arrangement of differently charged atoms. Also, the structural features of various cavities in biomolecules are complicated. All these possibilities makes our method of single water entropy calculation more significant. Using our method, it is possible to measure the entropy of any water molecule present either in the interior or exterior of biomolecules. Thus, there exists numerous systems of biological significance where the method can be applied to calculate entropy of individual water molecules.

The structure and shape of biomolecules often alter the thermodynamics of water molecules. While inside a hydrophobic cavity, the water molecules are expected to be in an energetically unfavorable and entropically favorable condition. By altering the shape or hydrophobicity of such cavities, the thermodynamic properties of water can be modified. Hence as a future goal, the single water entropy method can be applied to water near surfaces with various shapes and interactions to identify the pattern of entropy with surfaces' geometric characteristics. These findings then can be extended to biomolecular cavities to obtain a general relation between water entropic behavior and shape of cavities.

The other application of single water entropy calculation can come from the study of water near differently charged surfaces that mimics the structures of biomolecules. The different patterns

of charged atoms can be expected to alter water behavior in its vicinity. Once again, a systematic study can be carried out to identify general water behavior near different surfaces.

While in this thesis, we have found new insights on the behavior of supercooled water, our study has been limited to only pure water and for water near ions. The fact that proteins and nucleic acids also show a change in their dynamic features at a similar temperature where water exhibit its fragile to strong crossover, the exact structural and thermodynamic features of water near these systems at such low temperatures has not been probed at individual molecule level. While any protein or peptide surface can reveal new information in this regard, the water near antifreeze proteins also present systems which can be explored in great details. Our investigation at single molecule level near such antifreeze proteins can help future studies to develop better materials to prevent ice formation.

8.3 References

1. Nettels, D.; Müller-Späth, S.; Küster, F.; Hofmann, H.; Haenni, D.; Rüegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (49), 20740.
2. Best, R. B.; Zheng, W.; Mittal, J. *J. Chem. Theory. Comput.* **2014,** *10* (11), 5113-5124.
3. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell Jr, A. D. *Nat. Methods* **2016,** *14*, 71.
4. Robustelli, P.; Piana, S.; Shaw, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **2018,** *115* (21), E4758.
5. Bashardanesh, Z.; van der Spoel, D. *J. Phys. Chem. B* **2018,** *122* (33), 8018-8027.
6. Copeland, R. A. *Nat. Rev. Drug Discov.* **2016,** *15* (2), 87-95.
7. Bernetti, M.; Rosini, E.; Mollica, L.; Masetti, M.; Pollegioni, L.; Recanatini, M.; Cavalli, A. *J. Chem. Inf. Mod.* **2018,** *58* (11), 2255-2265.
8. Schuetz, D. A.; Bernetti, M.; Bertazzo, M.; Musil, D.; Eggenweiler, H.-M.; Recanatini, M.; Masetti, M.; Ecker, G. F.; Cavalli, A. *J. Chem. Inf. Mod.* **2018**.
9. Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H.-P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C. *J. Chem. Theory. Comput.* **2018,** *14* (7), 3859-3869.
10. Ganotra, G. K.; Wade, R. C. *ACS Med. Chem. Lett.* **2018,** *9* (11), 1134-1139.