SOLVATION ENERGY OF BIOMOLECULAR STRUCTURES: A STUDY OF THE EFFECT OF SALT ON BIOMOLECULES THROUGH IMPLICIT AS WELL AS EXPLICIT SOLVATION METHODS

Mihir Date
Clemson University, mihirdate11@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations
Part of the Chemistry Commons

Recommended Citation
Date, Mihir, "SOLVATION ENERGY OF BIOMOLECULAR STRUCTURES: A STUDY OF THE EFFECT OF SALT ON BIOMOLECULES THROUGH IMPLICIT AS WELL AS EXPLICIT SOLVATION METHODS" (2012). All Dissertations. 1019.
https://tigerprints.clemson.edu/all_dissertations/1019

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
SOLVATION ENERGY OF BIOMOLECULAR STRUCTURES: A STUDY OF THE EFFECT OF SALT ON BIOMOLECULES THROUGH IMPLICIT AS WELL AS EXPLICIT SOLVATION METHODS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Chemistry

by
Mihir Shashikant Date
December 2012

Accepted by:
Dr. Brian N. Dominy, Committee Chair
Dr. Steven J. Stuart
Dr. Jason D. McNeill
Dr. Emil G. Alexov
ABSTRACT

In the current dissertation, studies related to solvation energy of protein structures using implicit as well explicit solvation methods have been discussed. Special focus is given to explore effect of salt on the fold stability of proteins and enzymes. Salt plays a crucial role in the functioning of all proteins, enzymes and nucleic acids. Change in salt concentration of the medium has large impact on stability and activity of these biological macromolecules. Therefore exploring mechanism of salt effect on them and development of an efficient model to calculate the salt effect has fundamental as well as practical importance in the field of sciences.

In chapter two the development of an implicit solvation model to calculate salt effect on the fold stability of proteins and enzymes is shown. In combination of standard Poisson-Boltzmann formalism to calculate polar solvation energy, newly developed microscopic surface tension parameter as a function of ionic strength is used in the non-polar component of solvation free energy. The model was tested on series of Cold shock proteins whose stability as a function of NaCl concentration was calculated previously through experiments. Then the model was successfully used to explain the basis of experimentally observed increased stability of HIV-1 protease in the presence of high concentration of NaCl. Further, the same model also showed ability to capture salt specific Hofmeister effect on Cold shock proteins by using salt specific surface tension parameter.

In the third chapter, similar studies were extended through molecular dynamics simulations of explicit solvated aqueous systems of protein and salt. Effect of salt on the
translation and rotational motion of bulk water as well as water in different layers from protein surface was closely monitored. Self hydration of salt ions was seen to follow their rank in Hofmeister series. Alternatively effect of salt on rotational motion of water in different layers from protein surface showed that rank of an ion in Hofmeister series have no significant correlation with its effect on water structure making or breaking properties. The largest impact of salt on restricted motion of water was seen on the layer of water which is on the brink of being hydration water and bulk water. This is the same layer where water is been exchanged continually between hydrated water and bulk water. With these results, it can be articulated that effect of salt on the exchange rate of water between hydration shell and bulk may also be behind the origin of Hofmeister effect on protein.

After looking at the salt effect through explicit as well as implicit solvation methods, in chapter four we will compare generalized Born with a simple switching (GBSW) implicit solvent and explicit solvent using TIP3P water model effect of solvent viscosity on peptide dynamics. We compared both solvents to see if absence of solvent viscosity and equilibration of solvent’s degrees of freedom makes implicit solvent faster in sampling same conformational phase space than explicit solvent. To reach same equilibrium and sample phase space GBSW proved to be faster by factor of 10 than explicit solvent. An additional modified explicit solvent which thermodynamically identical to the original but higher in viscosity was studied too. The results confirmed that equilibrium properties of peptide calculated through implicit or explicit solvent matches and the efficiency of implicit solvent to sample similar phase space comes from inherent lack of friction and viscosity.
DEDICATION

To my parents Shashikant and Asha Date
ACKNOWLEDGMENT

I would like to thank my academic adviser Prof. Brian N. Dominy for his guidance and support throughout my Ph D career. He has always encouraged me and boosted my confidence to overcome obstacles in my research work. He taught me more than science, which will be a critical aid to prosper in future in countless number of ways.

I would also like to thank Prof. Steve Stuart for being an excellent teacher to me. I am highly benefited by his courses and discussions with him. I also thank my graduate committee members for providing important comments and suggestions that greatly improved this dissertation.

I am thankful to CCIT and personnel behind the Palmetto cluster. They were a key resource in my research.

Finally, I would like to thank my wife Chaitra, my brothers Rohit and Vinit, my parents Shashikant and Asha, and numerous other friends and relatives. Without their constant support and encouragement, this journey would not have been possible.

Mihir Shashikant Date

August 2012
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Implicit Solvent Models</td>
<td>2</td>
</tr>
</tbody>
</table>
Solvent Structure and Dynamics from Simulation with Explicit Solvent .......... 6
Hydration of Proteins ................................................................................. 8
Properties of Water near Protein Surface .................................................. 10
Biomolecule-ion Interactions ..................................................................... 11
Current Study .............................................................................................. 13
Summery .................................................................................................... 16

II. MODELING THE INFLUENCE OF SALT ON THE HYDROPHILIC
    EFFECT AND PROTEIN FOLD STABILITY ........................................... 17

Abstract .................................................................................................. 17
Introduction ............................................................................................... 19
Methods and Materials ........................................................................... 24
Results and Discussion ........................................................................... 29
Conclusions .............................................................................................. 49

III. STUDY OF THE EFFECT OF SALT ON WATER STRUCTURE AROUND
    PROTEIN THROUGH EXPLICIT SOLVATION METHOD. ..................... 51

Abstract .................................................................................................. 51
Introduction ............................................................................................... 53
Methods and Materials ........................................................................... 61
Table of Contents (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results and Discussion</td>
<td>65</td>
</tr>
<tr>
<td>Conclusions</td>
<td>88</td>
</tr>
<tr>
<td>IV. COMPARISON BETWEEN IMPLICIT AND EXPLICIT SOLVATION METHODS: EFFECT OF VISCOSITY ON SAMPLING PEPTIDE DYNAMICS</td>
<td>90</td>
</tr>
<tr>
<td>Abstract</td>
<td>90</td>
</tr>
<tr>
<td>Introduction</td>
<td>92</td>
</tr>
<tr>
<td>Methods and Materials</td>
<td>97</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>100</td>
</tr>
<tr>
<td>Conclusions</td>
<td>121</td>
</tr>
<tr>
<td>APPENDIXA</td>
<td>124</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>127</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The difference between the experimental data and two sets of calculated data using our model</td>
</tr>
<tr>
<td>3.1</td>
<td>Comparison of oxygen – oxygen (gr_OO) RDFs of water for different salt solution</td>
</tr>
<tr>
<td>3.2</td>
<td>Comparison of oxygen – hydrogen (gr_OH) RDFs of water for different salt solution</td>
</tr>
<tr>
<td>4.1</td>
<td>Structural elements in CSPA pdb id:1MJC</td>
</tr>
<tr>
<td>4.2</td>
<td>Diameter (width) of cross section area of different secondary structural elements, within protein as well as of isolated helices</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Salt effect on stability of CSP</td>
<td>34</td>
</tr>
<tr>
<td>2.2 Comparing experimental and calculated salt effect data</td>
<td>37</td>
</tr>
<tr>
<td>2.3 Free energy of unfolding of HIV-PR as a function of ionic strength</td>
<td>42</td>
</tr>
<tr>
<td>2.4 Salting in and salting out effect of salts A. Na$_2$SO$_4$ and B. NaSCN</td>
<td>46</td>
</tr>
<tr>
<td>3.1 Radial distribution functions (RDFs) calculated for TIP3P water oxygen-oxygen gr_OO, oxygen – hydrogen gr_OH and hydrogen – hydrogen gr_HH interactions in different salt solutions at different concentrations</td>
<td>66</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Radial distribution functions (RDFs) calculated for cation – TIP3P water oxygen interactions ( g(r) _ \text{cation-water}(O) ) (top row), hydration number of cations (middle row) and reorientational autocorrelation function of water dipole (bottom row) in different salt solutions at different concentrations</td>
</tr>
<tr>
<td>3.3</td>
<td>Dipole reorientational autocorrelation function decay time of water in salt solution as function of salt concentration</td>
</tr>
<tr>
<td>3.4</td>
<td>Self diffusion coefficient of water in different systems</td>
</tr>
<tr>
<td>3.5</td>
<td>Self diffusion coefficient of cation in different systems</td>
</tr>
<tr>
<td>3.6</td>
<td>Average number of hydrogen bonds and average life time of hydrogen bonds per water molecule, for water in different system studied</td>
</tr>
<tr>
<td>3.7</td>
<td>Dipole reorientational autocorrelation function decay time of water in protein-salt-water solution</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Backbone $\alpha$ carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for poly-proline PPII helix in different solvents.</td>
<td>102</td>
</tr>
<tr>
<td>4.2</td>
<td>Backbone $\alpha$ carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for poly-alanine $\alpha$ helix in different solvents.</td>
<td>103</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison between Poly-alanine $\alpha$-helix (on left) and poly-proline type II helix (on right).</td>
<td>106</td>
</tr>
<tr>
<td>4.4</td>
<td>Ramachandran plots for PPII and $\alpha$-helix.</td>
<td>109</td>
</tr>
<tr>
<td>4.5</td>
<td>Backbone $\alpha$ carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for CSPA in different solvents.</td>
<td>113</td>
</tr>
<tr>
<td>4.6</td>
<td>Timescale at which solute shows convergence of RMSF plotted against comparative viscosity of solvent.</td>
<td>116</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7 Atom-atom correlation function for PPII and α-helix (ALA) and cold shock protein CSP (1MJC)</td>
<td>119</td>
</tr>
</tbody>
</table>

xiii
CHAPTER ONE

INTRODUCTION

A precise understanding of biomolecules is incomplete without a consideration of solvent environment. Solvent plays a crucial role in functioning of biological system both directly by taking part in biological processes and indirectly by stabilizing biologically active conformations of protein, enzymes and nucleic acids and their assemblies. Since the beginning of life on the earth, an aqueous milieu out of all habitable environments distinguishes water as the most important biomolecular solvent. Solute-solvent interactions profoundly affect the conformational sampling relevant to biological activity and protein folding [1], which relies on the presence of water and in particular the hydrophobic effect [2, 3]. Sufficient hydration is also necessary to maintain biologically active B-form of DNA [4, 5]. Solvent also plays a key role in complex formation [6-10], ligand recognition [11, 12], DNA sequence recognition by DNA binding proteins [13] and smoothing of energy landscape of solutes [14]. Further, solvent interactions in biological environment are modulated by presence of co-solutes such as ions and variety of small molecules [15] and sometimes non-aqueous environments in form of lipid bilayers [16], where hydrophobic biomolecules can be sequestered from the aqueous environment.
Implicit solvation models:

For quantitative treatments of solute-solvent interactions Ben-Neim defined solvation process in a convenient way [17]. According to his definition the solvation of solute is defined as process of transferring a solute from a fix position in gas phase to a fix position in solvent at constant temperature, pressure and solvent composition. So the free energy of solvation ($\Delta G_{sol}$) can now be defined as reversible work required in the solvation process. This free energy of solvation incorporates both the free energy contributions related to direct solute-solvent interaction and those arising from internal changes in the solute and solvent upon solvation. Understanding these solute-solvent intermolecular interactions is important for any biomedical problem studied at molecular level. These intermolecular interactions control structure and function of biomolecules and therefore play an important role in the working of cellular system. Such interactions include several components such as contributions from linear, angular and torsion forces in covalent bonds, work required for cavitation, van der Walls forces and electrostatic forces. Generally solvation energy pertains to the last three types of interactions between solute and solvent viz. cavitation, dispersion-repulsion and electrostatics. Among these various components of molecular interactions, electrostatics has a special importance because of their long and short range and influence on polar or charged molecules which includes solute as well as solvent. Robust models for accounting electrostatic interactions are essential to understand the solvation properties of biomolecules and effect of solvation on biomolecular structure, function and activity. Therefore electrostatic forces are at the center in analysis of biological system at molecular scale. Within the frame
work of linear free-energy response theory, the electrostatic free energy contribution is one-half of the total solute-solvent electrostatic interaction energy. Free energy of solvation can be shown as

\[ \Delta G_{\text{sol}} = \Delta G_{\text{elec}} + (\Delta G_{\text{vdW}} + \Delta G_{\text{cav}}) \]  

(1.1)

Electrostatic and solvation models can be divided into two main classes. First is explicit solvent model, which treats the solvent in atomic detail and second is an implicit solvent model, which replaces the explicit solvent with dielectric continuum. While explicit solvent models offer some of the highest levels of detail they generally require extensive sampling to converge properties of interest. On the other hand, implicit solvent models trade details and some accuracy compared to explicit solvent for the pre-equilibration of solvent degrees of freedom by eliminating sampling of these degrees of freedom. Because of such pre-equilibration, implicit solvent models requires less computational efforts and have become popular in recent years. The very assumptions that give implicit solvation models their speed, namely pre-averaging of solvent and ions positions, are also their primary source of error and drawbacks. Implicit solvent models are capable of describing only non-specific interactions between solvent and solute. Explicit solvent methods needed to be used whenever detailed interactions between solvent and solute are important, such as solvent’s finite size effects on ion channels [18], strong and site specific interactions between solute and solvent or ions [19], and saturation of solvent polarization near membrane [20]. A third class of solvation model is hybrid implicit-explicit solvation models which take best of first two classes. In this method, region around solute up to certain radius contains explicitly represented water
molecules and beyond that region, a continuum dielectric exists. But these hybrid models are yet in development stage and are untested for wide range for problems [21].

With a few exceptions, implicit solvent models treat polar and non-polar interactions as separate processes. Continuum models for polar solvation generally attempt to solve or approximate solutions to partial differential equations for electrostatics in dielectric materials. Such approaches include Poisson-Boltzmann and Generalized Born treatments of biomolecular solvation [22, 23]. Accurate implicit models for non-polar treatment generally include two separate integrations for cavity formation term describing repulsive solute – solvent interactions and for an attractive term describing weak but favorable dispersion solute – solvent interactions. Along with the polar or electrostatic component of solvation free energy, Date & Dominy have showed that non-polar or cavitation component to be a dominant factor impacting stability of biomolecules, especially at high salt concentrations [24]. Cavity formation terms are popularly modeled by energies proportional to solvent accessible volume and surface area while attractive dispersive term generally require calculation of an integral or surface based integral approximations [25]. One popular approximation for the non-polar solvation free energy is a linear dependence between non-polar solvation free energy \( \Delta G_{cav} \) and the solvent accessible surface area model [26]

\[
G_{cav} = \gamma \cdot ASA
\]

(1.2)

where \( \gamma \) is surface tension which is typically chosen to reproduce the non-polar solvation free energy of alkanes [27] or model side chain analogues [28]. The surface tension value may assume same for all atom types or different values may be assigned for
different atom types. Although $(\gamma \cdot ASA)$ models have been successful and popular, they are also subject to several caveats, including widely varying choice of surface tension parameter [29] as well as inaccurate description of detailed aspects of non-polar energy [30], peptide conformations [31], and protein non-polar solvation forces [32]. Some of these problems have been solved by newer models that includes modified dispersion-repulsion interaction and solvent accessible volume terms [33] as well as different surface tension parameter we shown by us in chapter two.
Solvent structure and dynamics from simulations with explicit solvent

Interaction between solute and surrounding solvent and co-solutes affects thermodynamic as well as kinetic process of solute such as chemical reaction or conformation transition. For many enzymes, presence of water in active site is required to lower the energy barrier of transition state and to accelerate the reaction [34]. In contrast, less specific stochastic collisions between solute and solvent atoms provide both frictional drag forces and activation energy necessary to overcome energy barrier [35, 36]. The thermodynamic view of solvation can be connected to molecular level with the aid of statistical mechanics by using correlation functions that capture molecular organization of solvent and solute. Typically pair wise radial distribution functions (RDFs) $g_{ij}(r)$ are used to describe the density of distribution of an atomic or molecular species $i$ as a function of distance from another species $j$. It is possible to calculate thermodynamic quantities like total interaction energy $U_{ij}$ or potential of mean force (PMF) between solute and solvent using RDFs. These quantities can be calculates as,

$$U_{ij} = \frac{N\rho_i}{2} \int_0^\infty 4\pi r^2 g_{ij}(r)U(r)dr$$  \hspace{1cm} (1.3)$$

and

$$PMF = -RT \ln(g(r)) + c$$  \hspace{1cm} (1.4)$$

where $\rho_i$ is density of species $i$ and $U_i$ is potential energy function which is function of distance between $i$ and $j$.

Dynamic properties of solvent can typically be calculated through time correlation function. Because solvent dynamics might be substantially altered in close vicinity of solute due to solute-solvent interaction, such analysis often considers
conditional correlation functions that take solute proximity into account. A commonly
calculated property in this manner is distribution of solvent residence time in a specific
region according to correlation function [37-39]

\[ C_\alpha(t) = \frac{1}{N_{\text{solv}}} \sum_{i=1}^{N_{\text{solv}}} \frac{1}{N_{\alpha,i}} \sum_{t' = 0}^{t_{\text{max}}} p_{\alpha,i}(t', t' + t) \]  

(1.5)

where \( p_{\alpha,i}(t', t' + t) \) is binary function with a value of 1 if solvent molecule \( i \)
remains within confined area \( \alpha \) from time \( t' \) to \( t' + t \) and a value 0 otherwise. \( N_{\alpha,i} \) is the
number of times that a solvent molecule is present in \( \alpha \) and the outer sum is over all
molecules \( i \). Other correlation functions can be calculated to determine rotational
correlation times, the lifetimes of specific interactions such as hydrogen bond or ion-
peptide, and so on. In all cases the resulting correlation functions can be fitted to single,
double, or stretched exponential function [39, 40] to extract characteristic time scales that
can be compared with experimental measurements. Self diffusion coefficient is another
property easily modulated by solvent environment. It measures the rate of mean square
displacement at long timescales. It is calculated according to Einstein’s relation [41] from

\[ D = \frac{1}{6} \lim_{t \to \infty} \frac{d}{dt} \langle (r_i(t_0 + t) - r_i(t_0))^2 \rangle \]  

(1.6)

where the average is taken over all solvent molecules \( i \) and all time origins \( t_0 \).
Hydration of Proteins

Numerous theoretical and experimental studies of protein-water interactions have painted a complex picture of protein or peptide hydration that exceeds a range of temporal and special scales. The overall effect of aqueous solvent on proteins in terms of the hydrophobic effect and dielectric screening of electrostatic interactions between polar and charged groups still have some gaps and is away from a complete knowledge. For long peptides and proteins, solvent promotes the formation of regular secondary structure elements and folding into its native structure [42]. In short peptides, where hydrophobic residues cannot be fully sequestered from aqueous solvent environment, the electrostatic effect is dominant and results in competition of intermolecular salt bridges and hydrogen bonds [43]. In short peptides formation of single backbone hydrogen bond and individual solvation of carbonyl and amide groups, are energetically opposite in sign and equal in magnitude. The result is only slight enthalpic advantage of about 1 kcal/mol per residue, which is countered by an entropic cost [44]. This situation gives rise to a delicate balance of energy that allows short peptides to adopt a variety of conformations with similar relative free energy depending on sequence [1, 45]. Recent experimental and theoretical evidences suggests a general preference of short peptides for polyproline II (PPII) conformations with backbone dihedral torsion angles in the vicinity of $\Phi = -85^0$ and $\Psi = 140^0$ that is apparent in short alanine based peptides [46-48] and also found for longer peptides in unfolded state [49]. It is clear that solvent plays a significant role in stabilizing the PPII conformation, especially in polyalanine peptides, but the fact that exact mechanism remains subject of debate [50-52] is testament to complexity to the
peptide-water interactions even for the formation of α helices [53-55]. α helices are frequently observed in for alanine rich peptides [56] and are believed to be stabilized partly by helical dipole formation due to alignment of multiple backbone hydrogen bonds [57, 58] and partly by secondary effects due to side chain – water interactions [59]. Extended structures would have been favored due to entropic gain over solute’s part, but the more extensive backbone hydrogen bonding interactions with solvent incur, a penalty in solvent entropy is due to ordering of solvent molecules [48].
Properties of Water near protein surface

On the molecular level, interactions of surrounding water and protein are complex in nature. Water forms distinct hydration layer around protein surfaces that consists of one or more solvent shells containing water molecules with physical characters that differ from those in bulk. The first hydration layer extends about 3Å from protein surface and has an average density increased by about 10-15% over the value for bulk solvent [60, 61]. A small fraction of first hydration shell waters, on the order of 5-30%, exhibit reduced rotational and translational diffusion times and which is on the order of tens of picoseconds [62] and may show local interactions with solute [63]. The remaining water molecules retain ‘bulk-like’ properties, with diffusion times reduced by no more than 2-3 times than bulk solvent despite close interactions with solute [64]. With this, the modern view of hydration does not assume an ice-like water layer but allows for fluid-like individual water molecules that preferentially visit certain sites on protein surface where they exhibit temporarily retarded dynamics while continuing to exchange with surrounding bulk on sub-nanosecond timescales. Further, the notion of reduced solvent dynamics near protein surface needs to be differentiated based on observation that water molecules in first hydration shell largely retain bulk-like rotational dynamics [65] and lateral diffusion dynamics while most dramatically slowed down property is diffusion component perpendicular to protein surface [66]. While the influence of co-solute like salt ions on bulk water structure is most studied area, examining the same effect on water near protein surface is another important aspect of the same area that can be studied with molecular dynamics simulations.
Biomolecule-ion interactions

Within cellular environment biomolecules are immersed in aqueous media which contain ions and numerous other small molecules and macromolecules. Ions are generally classified into two classes which are kosmotropes or solvent structure makers and chaotropes or solvent structure breakers. These ions influence biomolecular processes and interactions in several different ways. It includes electrostatic screening, site specific ionic binding and preferential hydration effects. Electrostatic screening is a phenomenon in which the strength of electrostatic interactions within and between biomolecules is reduced by presence of salt ions. This is non-specific ionic effect and is described well, at low salt charge and concentration, by the Debye-Hückel theory [67] and related implicit solvation models like GB or PB. In site specific binding, ions interact with biomolecules by binding to specific sites in a manner similar to ligand binding [68]. Preferential hydration or Hofmeister effects are ion specific which arises due to competition between ions and water for binding to nonspecific site on biomolecules [69]. This competition is between weak protein-water and protein-ion interactions and therefore observed only at high salt concentration [70]. Kosmotropic ions are strongly hydrated and they decreases solubility of hydrophobic solutes by excluding them from solution phase as an aggregates (salting out). Chaotropic ions are weakly hydrated and tend to break down hydrogen bonding network of surrounding water molecules. These ions are themselves excluded from solution phase and pushed towards hydrophobic interface like protein surface, thereby interacting with protein side chains, they destabilizes protein in folded form (salt in) and tend to solubalize unfolded protein. A similar effect involves competition
between ionic species and water at protein-water interface which results in modulation of surface tension and thus can be easily incorporated in implicit models of solvation to account for non-polar term of solvation free energy [71]. Several hypotheses have been offered to explain the effect of salt ions on protein stability. The principle hypothesis states that stabilizer and destabilizer ions act indirectly by altering water structure and dynamics, but other hypothesis suggests that this is not the determining factor and the other factors such as excluded volume, affinity for protein surface, and ability to attenuate or accentuate the hydrophobic effect should be considered [72]. These hypotheses can be tested by examining hydrogen bonding network of bulk and hydration water and protein side chain dynamics to check influence of salt on it. The same phenomenon is looked into by taking examples of small organic solute molecules in aqueous solution with dissolved salts. Effect of salt on these solutes as well as water structure around it was examined [73]. Accounting for folding free energy of energetic changes of protein upon influence of salt would be of much interest since it will give an idea of the unfolding-folding equilibrium. Studies like the one mentioned just now would pertain more towards the unfolded state of the protein where small organic solutes can best mimic exposed amino acid side chains. The effect of salt on the other side of equilibrium that is folded state should be needful to complete the spectrum.
Current study

Starting with chapter two, we show our development of an implicit solvation model to access effect of ionic strength on protein fold stability. The model uses traditional Poisson-Boltzmann equation to calculate electrostatic solvation free energy of unfolding as a function of ionic strength. To calculate non-polar or hydrophobic contribution to protein fold stability, we use cavity model with newly developed microscopic surface tension parameter which is a function of ionic strength. The model was tested for its precision and accuracy by comparing calculated results using this model with experimental unfolding free energy data of thermophilic and mesophilic Cold shock protein B wild types and 27 of their mutants. The model proved to be in agreement with experimental results. As expected, the electrostatic component of solvation free energy calculated through Poisson-Boltzmann equation is adequate for showing correct qualitative trends and shows dominance in lower ionic strength region typically below 0.4 mol/L. Above this ionic strength the electrostatic component of solvation free energy alone, fails in bringing quantitative agreement compared to experimental results. We have shown that hydrophobic or non-polar component becomes dominant as the ionic strength is increased and our model proves to be important in having quantitative match with experiment at especially at high ionic strength. Later the model was successfully applied to HIV-1 protease to explain it experimentally observed increased stability and activity at high (1M) NaCl concentration and provide plausible mechanism of its stabilization by presence of salt.
Dissolved salt in water increases the surface tension on the interface. This effect of salt on protein fold stability can be studied by incorporating surface tension parameter in implicit solvent model which readers will see in chapter two. Salts are thought to originate their effect also by modulating the solvation water structure around proteins. Further to explore mechanism of Hofmeister effects on proteins, in the third chapter we examine effect of cation on structure and dynamics of water molecules in bulk and around protein surface. Systems of protein immersed in a box of TIP3P water and salt ions were studied via molecular dynamics simulations carried in isothermal and isobaric ensemble at 300K. Cations on the opposite extremes of Hofmeister series which are Mg$^{2+}$ and Cs$^+$, along with Na$^+$ which is more neutral and in the middle of series, and thermophilic (net charge –2) and mesophilic (net charge –6) cold shock protein were chosen for the study. The characteristic self diffusion coefficient of water and cation, orientation autocorrelation of dipole vector of water in different shells from protein surface and hydration number for protein, water and cation as well as average lifetime of hydrogen bond were calculated for all cations and thermophilic as well as mesophilic proteins. The results reveal structure of water around protein in terms of rotational motion does not show significant dependence on added salt but translation motion of water is affect by presence of salt and is ion specific. These findings throw light on the origin of Hofmeister effects on globular protein surface may be via direct interaction or surface tension modulation at the protein-water interface and not through altering structure of bound water to protein surface.
After inspecting the effect of salt on protein structure and dynamics by means of implicit solvation model and molecular dynamics simulations with explicitly solvated system, we complement this study by scrutinizing effect of type of solvation model on dynamics and kinetics of protein motions. We compare generalized Born with a simple switching (GBSW) solvation implicit model with explicit solvent using TIP3P water model. Using molecular dynamics simulations isolated standard secondary structures like polyproline II and α helices, and a β – barrel Cold shock protein A derived from *E. Coli* were simulated in implicit as well as explicit solvent. Normal mode analysis was used to confirm the motion of peptide structures. Average solvation free energy, RMSD, backbone RMSF and dihedral RMSD calculated at different time scales. Our results show significant dependence of dynamic properties and sampling speed on type of solvent model used. Protein motions in implicit solvent can be captured around ten times faster than explicit solvent. On the other hand equilibrium properties do not show significant dependence on the type of solvent used. To confirm our findings we extended the study with molecular dynamics simulations with additional explicit solvent, based on TIP3P water model and differing only in viscosity increased by factor of 10 (TIP3P-100). This was achieved by increasing masses of all the atoms of water by factor of 100. Our results with TIP3P-100 confirmed that enhanced sampling efficiency and speed of GBSW was originating from the lack of viscosity and stochastic friction offered by motion of solvent molecules compared with TIP3P.
Summary

In summary, the research presented in current dissertation provides molecular level insights into the understanding of effect of salt and solvent model on structure and dynamics of protein and water around its surface. Considering current knowledge of origin and mechanism of effect of salt on protein fold stability our study not only provides a new perceptive, but also has implications in development of solvent models for better efficiency and accurately accounting natural forces on protein. Detailed discussion about the research has been provided in following respective chapters.
CHAPTER TWO

MODELING THE INFLUENCE OF SALT ON THE HYDROPHOBIC EFFECT AND PROTEIN FOLD STABILITY

This work has been published as:


Accepted October 2011, Available Online June 2012

Abstract:

Salt influences protein stability through electrostatic mechanisms as well as through nonpolar Hofmeister effects. In the present work, a continuum solvation based model is developed to explore the impact of salt on protein stability. This model relies on a traditional Poisson-Boltzmann (PB) term to describe the polar or electrostatic effects of salt, and a surface area dependent term containing a salt concentration dependent microscopic surface tension function to capture the non-polar Hofmeister effects. The model is first validated against a series of cold-shock protein variants whose salt-dependent protein fold stability profiles have been previously determined experimentally. The approach is then applied to HIV-1 protease in order to explain an experimentally observed enhancement in stability and activity at high (1M) NaCl concentration. The inclusion of the salt-dependent non-polar term brings the model into quantitative
agreement with experiment, and provides the basis for further studies into the impact of ionic strength on protein structure, function, and evolution.
Introduction:

It has long been understood that salts have a significant impact on the stability and activity of proteins and nucleic acids, which constitute a foundation underlying cellular function. The intracellular and extracellular salt concentration varies with the organism and environment, but fluctuates typically in the range of 100-200 mmol/L for organisms living within physiological conditions consistent with mesophiles. Salts can significantly influence the stability of biomolecules by screening electrostatic interactions. For example, the repulsion between negative charges on the phosphate backbone of nucleic acids is screened by salt, contributing to the stability of biologically relevant DNA and RNA conformations [74]. The effect of salts on proteins has been demonstrated using numerous systems including multimeric complexes that have been shown to disintegrate into separate monomers upon changes in the environmental salt concentration. Some proteins within halophilic organisms have even adapted to function specifically within high salinity environments, destabilizing under physiological salt concentrations consistent with mesophilic environments [75, 76]. The development of accurate physical models describing the thermodynamic impact of salts on macromolecules could lead to a broader understanding of biomolecular structure and function.

Within a cellular setting, biomolecules are solvated in an aqueous environment containing salt ions and numerous other solutes. The influence of ions on biomolecular interactions is mediated through electrostatic screening, site-specific binding, and preferential hydrophobic effects or Hofmeister effects. The role of salt concentration in protein stability can be determined by measuring the unfolding transition as a function of
Some ions stabilize proteins by binding to specific sites. This ligand-induced ion-specific stabilization is usually observed below 0.2 mol/L ionic strength [77]. Bulk ionic strength results in the screening of surface charge-charge interactions primarily at lower salt concentrations. Hofmeister effects, which are dominant at higher salt concentrations, strengthen the hydrophobic effect by increasing the surface tension of the solvent, or by stabilizing peptide dipoles through specific ionic interactions [78]. Theoretical modeling can provide a basis from which the different mechanisms associated with salt effects may be assessed and compared.

The electrostatic screening effect (and the effect on self polarization energies), primarily related to bulk ionic strength, can be studied through continuum electrostatic models such as those based on a Poisson-Boltzmann (PB) formalism [79, 80]. The PB equation can be used to describe the electrostatic potential from the reaction field of a system containing a solute with a fixed charge distribution and a surrounding mobile charge distribution representing the salt. Mobile charges are modeled by a Boltzmann distribution with respect to the electrostatic potential generated by the fixed charges of the solute [22].

$$\nabla \cdot \left[ \epsilon(\vec{r}) \nabla \varphi(\vec{r}) \right] = -4\pi p^f(\vec{r}) - 4\pi \sum_{i} C_i^{\infty} Z_i \lambda(\vec{r}) \cdot e^{-Z_i \varphi(\vec{r})/\kappa_0 T}$$  \quad (2.1)

Equation 1 is the Poisson-Boltzmann equation, which describes electrostatic interactions between solute and solvent molecules where $p^f$ is charge density of fixed charges. $C_i^{\infty}$ is the concentration of ion $i$ and $Z_i$ is the charge of the ion. The second term on right hand side represents mobile charges typically restricted to the solvent region.
This term contains the ionic strength and charge density associated with the mobile charges. The distribution of mobile charges around solute’s fixed charges is modeled by the Boltzmann factor \( e^{-Z_i q \phi(r)/\kappa B T} \). This special case of Debye-Hückel theory with a 1:1 electrolyte accounts for continuum solvent polarization effects involving a salt distribution around the solute [81]. Using this model, the electrostatic free energy of solvation can be determined, as well as the role of ionic strength on electrostatic screening.

While PB (or generalized Born [82]) models are used to calculate polar component of the solvation free energy and the role of ionic strength, surface area based models are popular in accounting for the non-polar hydrophobic component. These models express the non-polar solvation term as linearly proportional to the solute’s surface area. The most famous being the cavity model developed by Sitkoff et. al., which is based on partition coefficients for a series of hydrocarbons [83].

\[
\Delta G_{\text{non-polar}} = \gamma \times ASA
\]  

(2.2)

The cavity model is based on the scaled practical theory and expresses the non-polar solvation energy as a product of the solute’s solvent accessible surface area (ASA) and surface tension. This proportionality constant between non-polar solvation free energy and solute’s accessible surface area is described in the current study as the microscopic solute-solvent surface tension \( \gamma \) [27]. Recently, alternative models describing the hydrophobic effect have been described in the literature. Some alternative models involve the inclusion of dispersion integrals and solvent accessible volume terms [84, 85]. They indicate that the work of cavitation for a solute should depend on its
solvent accessible surface area (SASA) and solvent accessible volume (SAV), with the SASA term dominating for large solutes. Another popular model to account for non-polar solute-solvent interactions is the dewetting model by Berne and co-workers [86]. This model accounts for non-polar forces acting on a surface through the critical distance between surfaces in contact, the contact angle and surface tension of water.

While both of the models provide alternatives to the cavity model to estimate the non-polar component of solvation free energy, the cavity model has been validated numerous times in the literature especially for application to biomolecular system. In addition, its simplicity and current wide-spread use suggested this model for the current study. Parameterized for aqueous solutions, the cavity model provides a quantitative description of the hydrophobic effect at a molecular scale. The model is based on the macroscopic description of surface free energy or surface tension. The solute solvent microscopic interface on the surface of biological macromolecules is far different than an oil-water or water-air macroscopic interface. The cavity model therefore requires the predetermination of a microscopic surface tension parameter that should also be dependent on the environmental salt concentration.

Here we develop a model to describe the effect of ionic strength on protein fold stability involving traditional PB continuum electrostatic theory and a surface area based hydrophobic term. The cavity model used here includes a microscopic surface tension that is now described as a function of ionic strength. We validate our model by calculating the effect of ionic strength on the stability of a family of Cold Shock Proteins (CSP’s). Further we successfully apply our model toward the HIV-1 protease enzyme
(HIV-PR) to discover the mechanisms underlying the experimentally observed influence of NaCl on the enzyme’s stability and activity. The results from these calculations are found to be in close agreement with corresponding experimental results [87, 88].
Methods and materials:

1 Protein and enzyme models:

We have studied the effect of salt on the fold stability of the cold shock protein family derived from the mesophilic bacterium *Bacillus subtilis* (Bs-CspB) and the thermophilic bacterium *Bacillus caldolyticus* (Bc-Csp). We used the wild type and mutants from these proteins to authenticate our model describing the effect of salt through both electrostatic and hydrophobic contributions to the free energy of solvation. The crystal structures of the cold shock proteins, 1CSP and 1C9O, were used as a starting point [89, 90]. Missing heavy atoms and hydrogens were built and a gentle minimization technique, involving a stepwise reduction in a harmonic force constraint on the positions of atoms available in the crystal structure, was applied to the CSP crystal structures using the CHARMm (c32b1) molecular mechanics package [91]. All minimized structures yielded an RMSD below 0.7 Å relative to the corresponding crystal structures. Mutants of these proteins were modeled using MODELLER software package [92]. The mutants modeled and examined in this study are E3R, E3L, A46E, E66L, E3R/E66L, E3R/T64V/E66L for Bs-CspB and L66E, Q53E, E46A, T31S, S24D, G23Q, E21A, Y15F, N11S, R3E, R3L, R3A, Q2L, 3E/E46A/L66E, G23Q/S24D, R3E/L66E, R3E/E21A, R3E/E46A, E46A/L66E for Bc-Csp. Mutated residues were minimized in the context of a fixed protein environment in order to remove vdW clashes and improve electrostatic interactions.

The presently available structures of apo form HIV-PR include different types of conformations including semi-open, open and curled. Wild type apo form structures of
the semi-open and curled conformation are available. To investigate the increased 
stability of HIV-PR with increased NaCl concentration, two crystal structures of HIV-PR 
which are 1HHP (semi-open) and 3PHV (curled) from the protein databank were used 
[93, 94]. HIV-PR is only active as a homo-dimer and has an aspartate residue at the 25\textsuperscript{th} 
position in each monomer, in its active site, that is essential to its catalytic activity. In the 
enzyme’s dimeric form two such aspartate, one from each monomer, come together to 
form the active site. Numerous theoretical and experimental studies have been performed 
in the past to determine protonation state of these aspartates in the active site of HIV-PR 
dimer [95, 96]. Work by Smith \textit{et. al.} on HIV-PR, through C\textsuperscript{13} NMR experiments, 
suggests that HIV-PR in its unbound or apo state shows two chemically equivalent 
aspartic side chains in the catalytic site and they are ionic [97]. Hence, we modeled these 
catalytic aspartates in their ionic state of side chains. Coordinates of the structures for 
HIV-PR were taken from RCSB Protein Data Bank. Any missing atoms in the crystal 
structures were added and then minimized using the same strategy applied to the CSP 
wild type structures within the CHARMM (c32b1) molecular mechanics package.

2 Calculating the “Salt effect”: The polar or electrostatic component:

As described earlier, the solvation free energy is composed of two components: 
the electrostatic or polar and the hydrophobic or non-polar components.

\[
\Delta G_{solvation} = \Delta G_{elec} + \Delta G_{non-polar}
\] 

(2.3)
Using the PB model, the electrostatic component of the unfolding free energy \((\Delta G_{el}(I))\) can be calculated as a difference between the unfolded and folded states of a protein.

\[
\Delta G_{el}(I) = G_{el}^u(I) - G_{el}^f(I)
\]  

(2.4)

where \(G_{el}^u(I)\) is the electrostatic free energy of unfolded state and \(G_{el}^f(I)\) is the electrostatic free energy of folded state at some ionic strength \((I)\). The crystal structures were used as the folded state. The unfolded state of a protein is the point where tertiary and most secondary structure is lost. It is modeled in this study by assuming each residue (i) interacts only with its two sequential neighbors on each side (i-2, i-1, i+1, i+2). The local conformation of this five residue segment representing the unfolded state is kept unaltered as it appears in the folded state. A properly normalized sum, over all the residues, of the self and interaction (Coulombic and cross-polarization) energies of each residue in the middle of the segment, preventing the double counting of interaction energies, is the total electrostatic free energy of unfolded state. Thus, the electrostatic free energy of the unfolded state becomes,

\[
G_{el}^u(I) = \sum_1^N G_{el}^u(I)
\]  

(2.5)

where \(N\) is total number of residues in the protein. The same unfolded state model was used in an earlier study where it was also compared to alternative models of the unfolded state[98]. In this study, the electrostatic component of salt effect was shown to be robust towards the choice of unfolded state model. Structure-based thermodynamic analyses of HIV-PR indicate the folding of individual monomers occurs prior to the dimerization of monomers, resulting in the final and active conformer [99, 100]. Consequently, in the
case of HIV-PR, in addition to the unfolding free energy of monomers shown by equation 5, the free energy contribution from the dissociation of the dimmer was also included to account for total free energy of unfolding. It can be shown as,

\[ G_{el}^{F-D}(I) = G_{el}^{A}(I) + G_{el}^{B}(I) \]  \hspace{1cm} (2.6)

where \( G_{el}^{A}(I) \) and \( G_{el}^{B}(I) \) are the electrostatic free energies of folded monomer A and monomer B and \( G_{el}^{F-D}(I) \) is energy of folded dimer of HIV-PR. The two contributions from dimer dissociation and monomer unfolding were simply added together to account for total unfolding free energy of HIV-PR. The salt dependence of the unfolding free energy, or the “salt effect” \( \Delta \Delta G_{el}(I) \), is the difference in the electrostatic component of unfolding free energy calculated at some ionic strength and reference ionic strength \( (I') \).

\[ \Delta \Delta G_{el}(I) = \Delta G_{el}(I) - \Delta G_{el}(I') \]  \hspace{1cm} (2.7)

The model used to calculate the electrostatic free energy of unfolding involved a finite difference solution of the linearized Poisson-Boltzmann equation and a thermodynamic cycle that involved a structural model for folded and unfolded states[101]. Charged states of titratable residues for folded as well as unfolded state models of all CSP and HIV-PR structures, were assigned based on their isolated states at neutral pH. It is neither intended nor within the scope of this study to explicitly incorporate the effect of pH while investigating the salt effect on protein fold stability. In order to see the effect of surrounding pH along with this salt effect study, an appropriate methodology can be utilized to first calculate the pKa’s and corresponding charged states of titratable residues in both the folded and unfolded conformations at each solution ionic strength. Poisson-Boltzmann calculations[22] were performed within the CHARMM
(c32b1) molecular mechanics package. Chosen grid parameters included a grid point density of 0.5 Å⁻¹, and a grid size of 50 Å³ and 68 Å³ for CSP and HIV-PR respectively. A variety of different grid spacings were examined ranging from 1.0 to 0.2 Å⁻¹ and the polar salt effect was found to converge at a grid density of 0.5 Å⁻¹. Atomic charges and vdw radii were assigned according to CHARMM param27 all hydrogen force field. The vdw molecular surface was chosen for electrostatic as well as hydrophobic term calculation.

The hydrophobic solvation term was simply added to the component related to the electrostatic solvation free energy to provide a model of the salting out contribution to protein stability with increasing salt effect. The protein dielectric constant is not a universal constant but a parameter that depends on the model used [102]. Theoretical studies on different proteins and enzymes have been performed with protein interior dielectric constants varying from 2 to 20 [79, 103-106]. Salt effect studies were conducted on CSP as well as HIV-PR using protein interior dielectrics of 2, 4, 10 and 20 (data not shown). While all of them recapitulated the qualitative results, a dielectric of 4 resulted in the best quantitative agreement with experimental data. Consequently, a protein interior dielectric constant of 4 was used in modeling both CSP and HIV-PR. This value for the interior protein dielectric constant has also been used previously and was shown to yield accurate results [98, 106]. The dielectric constant for the surrounding continuum solvent water was assigned a value of 80.
Results and Discussion:

1. The non-polar solvation energy as a function of ionic strength:

We calculated the change in stability of CSP with increasing salt concentration known as the salt effect. In implicit solvation models, the influence of increasing salt concentration on protein stability is described in mainly two ways. The first is through electrostatic influences on self-polarization energy and cross-polarization energy, or screening of electrostatic interactions. This effect is predominant for all types of salt typically below a concentration of 0.5 mol/L. The second is through the surface tension defined at the solute-solvent interface, which increases with an increase in the salt concentration. In an aqueous solvent, this increase in the surface tension increases the hydrophobic effect [107-109] (Eq. 2.2). Consequently to calculate the true salt effect on protein stability, it is important to take into account the hydrophobic effect along with the electrostatic contribution to the unfolding free energy [83]. Our model uses a cavity model to describe the non-polar or hydrophobic component of solvation (Eq. 2.2).

The relationship between the microscopic surface tension and the well-known macroscopic analog provides the basis for a more generalized theory of non-polar solvation and the corresponding hydrophobic effect. The free energy associated with the transfer of a non-polar solute from oil into water (microscopic) and the interfacial free energies between non-polar liquids and water (macroscopic) provide alternate measures of the free energy per unit area or the ‘surface tension’. Based on published data, these two measures of surface free energy are mutually inconsistent [110]. While macroscopic surface tension can be measured directly through experiments, microscopic surface
tension is estimated through the partition coefficients and oil-water transfer free energy of solutes. Shown through solvent transfer experiments for wide range of hydrocarbon molecules, hydration free energy depends linearly on the burial of solvent accessible surface area. Based on experimentally determined transfer free energies, a microscopic surface tension value \( \gamma_{\text{micro}} \) was determined. It ranges from 25 cal/Å\(^2\) to 31 cal/Å\(^2\) for different groups of alkanes including linear (28 ± 2 cal/Å\(^2\)), branched (31 cal/Å\(^2\)) as well as cyclic and aromatic (25 cal/Å\(^2\)) [26, 27, 111]. These groups best mimic amino acid side chains in proteins and enzymes. Thus the microscopic surface tension parameter derived from hydration free energies through solvent transfer experiments of these groups can be used towards the evaluation of the non-polar or hydrophobic component of the solvation free energy of proteins. We take the average microscopic surface tension from the different classes of hydrocarbon solutes mentioned above. The average value for \( \gamma_{\text{micro}} \) comes to 28 cal/Å\(^2\), which is almost two-fifths of the macroscopic oil-water interface surface tension value \( \gamma_{\text{macro}} \) of ~ 72 cal/Å\(^2\) [17, 111, 112]. Selection of this value is consistent with the type of surface area used in PB electrostatic calculations.

To develop a model of the microscopic surface tension as a function of ionic strength, we begin with the macroscopic oil-water interface surface tension which is approximately 72 cal/Å\(^2\) or 0.3 kJ/Å\(^2\). The surface tension of the electrolyte solution increases with increase in ionic strength. Under isobaric and isothermal conditions, the dependence of the surface tension on the electrolyte concentration can be given from Gibbs adsorption equation [113].

\[
d\gamma = -\sum_i \Gamma_i d\mu_i \tag{2.8}
\]
where, $\Gamma_i$ is surface excess/deficiency of ion I, and $\mu_i$ is the chemical potential of the salt ions at the interface. In this relation both of the parameters $\mu_i$ and $\Gamma_i$ are dependent of the concentration of salt ions in the solution. Equation 2.8 describes the dependence between the surface tension of an electrolyte solution and its ionic strength. In order to calculate the non-polar component of the solvation free energy as a function of ionic strength, it is essential to accurately calculate the increment in surface tension rather its absolute value.

The increment in macroscopic surface tension as a function of NaCl concentration $(\partial \gamma_{\text{macro}} / \partial l)$ is 1.64 dynes/cm or 9.9 J/Å²mol [113, 114]. This value varies with the specific salt used, but is similar for 1:1 salts such as NaCl or KCl. This derivative of the macroscopic surface tension remains constant up to a salt concentration of 1 mol/L. Similarly, the salt concentration should have a corresponding impact on the microscopic surface tension. We assume that the impact of the salt concentration on the microscopic surface tension will be reduced but still enables the ions to raise the surface energy at the interface and will be smaller by the same proportion between macro and microscopic surface tension. As pointed out before we consider the fact that surface of molecular size should respond in a manner similar to that of macroscopic surface to changes in the thermodynamic activity and hydrogen bonding properties of bulk water caused by adding a Hofmeister salt [80]. Consequently the changes in macroscopic surface tension produced by a Hofmeister salt should provide a useful guide to what is happening at the microscopic surfaces [115]. Based on the fact that the microscopic surface tension value $(\gamma_{\text{micro}})$ chosen in this study, is approximately 2/5 of the macroscopic surface tension.
value \((\gamma_{\text{macro}})\), we assume that the increment in microscopic surface tension with increasing ionic strength \((\partial \gamma_{\text{micro}} / \partial I)\) is \(2/5\) of the value associated with the macroscopic surface tension \((\partial \gamma_{\text{macro}} / \partial I)\). Thus, the partial derivative \((\partial \gamma_{\text{micro}} / \partial I)\) becomes,

\[
\frac{\partial \gamma_{\text{micro}}}{\partial I} = \frac{1}{2.6} \cdot \frac{\partial \gamma_{\text{macro}}}{\partial I}
\]

Equation 2.9 enables us to calculate increase in microscopic surface tension with ionic strength. As the macroscopic surface tension increases linearly with increasing ionic strength, we assume increment in microscopic surface tension will stay constant up to a salt concentration of 1 mol/L. Based on the facts and assumptions stated above, the salt concentration dependent microscopic surface tension becomes

\[
\gamma_{\text{micro}}(I) = \gamma_{\text{micro}}(0) + \frac{\partial \gamma_{\text{micro}}}{\partial I} \cdot I
\]

Using equation 9 and equation 10, it is possible to calculate the microscopic surface tension as a function of ionic strength. This can be straightforwardly used in conjunction with the cavity model to estimate the non-polar or hydrophobic component of the solvation free energy as a function of ionic strength.

2 Assessing the model: Fold stability of CSP variants as a function of ionic strength

After describing how salt influences electrostatic screening and the hydrophobic effect, we validated our model by calculating the salt effect on fold stability of the cold-shock protein (CSP) family and several associated mutants. CSP’s are small, monomeric
proteins expressed by mesophilic *Bacillus subtilis* (Bs-CspB) and thermophilic *Bacillus caldolyticus* (Bc-Csp). Within the CSP family, the mesophilic protein was reported to show an increasing stability while decreasing stability was reported for the thermophilic protein with increasing NaCl concentration [90]. It was found that the increased stability of Bs-CspB originates entirely from two residues, which are E3 and E66 in Bs-CspB while in the thermophile these positions correspond to R3 and L66 [116]. Salt screens the unfavorable pair wise Coulombic interaction between E3 and E66 in Bs-CspB, resulting in a halophilic response with increasing ionic strength. Experimentally, the unfolding free energies of 19 variants of Bc-Csp and 6 variants of Bs-CspB have been determined under a variety of ionic strength conditions [88]. The effect of ionic strength was studied previously but was restricted only to the electrostatic component of the free energy of solvation [98, 106].
Figure 2.1: Salt effect on stability of CSP. The salt effect is shown only for 4 representative structures among 27 total structures studied in this work. For each structure, experimental unfolding free energy as a function of NaCl concentration [88] are shown by dashed lines and calculated results are shown by solid lines in A as well as
B. Legends shown in B are same for A as well. A. Calculated stability includes only electrostatic contribution to protein stability using our theoretical model B. Calculated stability includes electrostatic contribution as well as hydrophobic term using our model. Only when the hydrophobic term is added to the electrostatic component of solvation free energy do the theoretical calculations agree with experimental results. The experiments were conducted in a buffer of 0.1mol/L ionic strength. To match this, our reference point for calculating the salt effect is 0.1mol/L and not zero. The error bars in B show the limits of the model based on the previously described range of microscopic surface tension parameters (between 25 and 31 cal/Å²mol).

The influence of ionic strength on the stability of the cold shock protein family is shown in figure 2.1. The solid lines in figure 2.1A show the salt effect involving only the electrostatic component determined through our calculations, while the solid lines in figure 2.1B show the complete salt effect involving both electrostatic and hydrophobic components as a function of NaCl concentration calculated through our model. Dashed lines figure 2.1A and B both, represent the experimentally determined salt effect [88]. These figures clearly demonstrate that the electrostatic influence of ionic strength toward fold stability alone is insufficient to reproduce the experimentally observed behavior. The electrostatic stability profile shows good qualitative agreement but is insufficient in providing quantitative agreement with experiment. The saturating effect of the electrostatic stability in figure 2.1A was seen as the ionic strength was increased above ~0.5 mol/L. This is not seen in the experimental results as experimental data represent a
combination of electrostatic as well as hydrophobic influences on protein stability. At higher salt concentrations, although the electrostatic effect saturates, increase in surface tension and the hydrophobic force continue to stabilize the protein fold. Inclusion of the ionic strength dependent microscopic surface tension through the cavity model accounts for the enhanced fold stability through the hydrophobic effect at higher salt concentrations, while not significantly impacting the predicted changes in fold stability at salt concentrations below 0.5 mol/L.

The result is a model with improved agreement with experimental studies over a broader range of ionic strength. This is seen from figure 2.1B where the theoretical data demonstrate excellent agreement with experimental data. The hydrophobic effect is dominant and the electrostatic screening effect of salt is saturated at high ionic strengths typically from 0.5 mol/L to 1.0 mol/L. Therefore in the same range of ionic strength, change in free energy of solvation as a function of ionic strength is coming almost entirely through hydrophobic effect. The average change in free energy as a function of ionic strength, based on experimental results, calculated for points between 0.7 mol/L to 1.0 mol/L ionic strength, comes out to be 4.13 (±1.3) kJ/mol. The corresponding average change in free energy based on our model over the same NaCl concentration range is 3.01 (± 1.1) kJ/mol. This close agreement between theory and experiment provides further evidence supporting the validity of microscopic surface tension function developed in this study.
Figure 2.2: Comparing experimental and calculated salt effect data. A. Calculated values include only electrostatic term. B. Calculated values include the electrostatic and hydrophobic terms. Data points represent 27 structures studied from 0.1 mol/L to 1.0 mol/L ionic strength giving 243 data points in total. The correlation between
experimental and calculated results improves drastically from A to B. Dotted lines show upper and lower limits of model using microscopic surface tension at extremes (25 and 31 cal/Å²mol).

After observing the qualitative and quantitative behavior of the salt effect model applied to four representative CSP structures, we applied the same model to the entire set of CSP structures to access the validity of our model. Figure 2.2A demonstrates the broader agreement between experimental and calculated data over a variety of protein variants and salt conditions. It shows a comparison between salt-dependent protein stabilities determined from experimental data and calculations using 27 mutant structures from the CSP family at ionic strengths ranging from 0.1 mol/L to 1.0 mol/L with 0.1 mol/L increment each step. The correlation between the electrostatic component and experimental unfolding free energies (figure 2.2A) is very poor ($R^2$=0.21), but increases dramatically ($R^2$=0.87) when the hydrophobic or salting out term is added to calculated electrostatic unfolding free energy (figure 2.2B).

<table>
<thead>
<tr>
<th></th>
<th>At 0.1 mol/L salt</th>
<th>At 1.0 mol/L salt</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \Delta G_{elec}$</td>
<td>-1.07</td>
<td>-5.73</td>
</tr>
<tr>
<td>$\Delta \Delta G_{elec} + \gamma \cdot ASA$</td>
<td>-0.61 ($\pm$ 0.05)</td>
<td>-1.14 ($\pm$ 0.4)</td>
</tr>
<tr>
<td>$\langle \Delta \Delta G_{hydrophobic} \rangle$</td>
<td>0.46</td>
<td>4.61</td>
</tr>
</tbody>
</table>

Table 2.1: The difference between the experimental data and two sets of calculated data using our model. $\Delta \Delta G_{elec}$ describes only the electrostatic component while
\( \Delta \Delta G_{elec} + \gamma \cdot ASA \) describes both the electrostatic and hydrophobic components of solvation free energy. \( \langle \Delta \Delta G_{hydrophobic} \rangle \) (in kJ/mol), is the average difference between \( \langle \Delta \Delta G \rangle \) for two sets of theoretical calculations. Numbers in parenthesis represents error estimation using microscopic surface tensions \( (\gamma_{\text{micro}}) \) at extremes (25 and 31 cal/Å²mol). \( \langle \Delta \Delta G_{hydrophobic} \rangle \) is around 10 fold larger going from 0.1 to 1.0 mol/L salt concentration demonstrating the importance of including the hydrophobic effect.

Table 2.1 summarizes data points from figure 2.2 only at 0.1 mol/L and 1.0 mol/L salt concentrations. It clarifies the difference between the experimental data set and two sets of theoretical calculations. It is seen that in comparison to the experimental results, at 0.1 mol/L theoretical calculations involving only electrostatic component under predict by only 1.07 kJ/mol. Relative to the electrostatic component, the hydrophobic influence is negligible at 0.1 mol/L. After adding the hydrophobic term, the agreement between the theory and experiment improves though slightly. The average difference \( \langle \Delta \Delta G_{hydrophobic} \rangle \) between two sets of calculations is very small \( \langle \Delta \Delta G_{hydrophobic} \rangle = 0.46 \text{ kJ/mol} \) demonstrating that the magnitude of the hydrophobic effect is minimal at low ionic strengths and salt primarily plays a role of screening electrostatic interactions at low ionic strength. However, at 1.0 mol/L ionic strength, the average error between two sets of calculations is much larger \( \langle \Delta \Delta G_{hydrophobic} \rangle = 4.61 \text{ kJ/mol} \). This demonstrates that at high ionic strength the hydrophobic or non-polar component becomes a significant and dominant mechanism underlying the effect of salt on protein stability.
Our model demonstrates very good agreement with experimental results; however a perfect match is not expected. Currently, our implicit model neglects specific ionic interactions and Hofmeister effects. These specific ionic interactions are negligible at low salt concentration but their influence increases with increasing salt concentration. In addition, the unfolded reference state chosen in our model is not necessarily representative of the true denatured state ensemble. Our unfolded state model for all CSP variants is same and presumes no long-range interactions. The experimental data reports thermally denatured proteins where protein unfolded state may not be extended random coils. Our unfolded state for CSP likely provides an upper bound on the true influence of salt on protein fold stability.

3 Fold stability of HIV-PR as a function of ionic strength:

Following the validation of our model on the CSP family, we then applied it toward characterizing the ionic strength dependent fold stability of HIV protease. Interestingly, though mesophilic in nature, HIV-PR has been shown to exhibit increased stability and catalytic activity with increasing salt concentration [87]. It was pointed out that the Michaelis constant $K_m$ values, integral to effective rate constants in enzymatic catalysis, are dependent on conditions like pH and salt concentrations [117]. These environmental factors can manifest their influence on catalytic rate through enzyme conformational stability. Through this study, we seek to determine the physical basis underlying the effect of ionic strength on the conformational stability of HIV-PR, and its relation to the activity of this enzyme.
In order to analyze the stability of HIV-PR, we first calculate the electrostatic contribution of the unfolding free energy with increasing ionic strength. From electrostatic calculations shown in figure 2.3, HIV-PR is destabilized with increasing ionic strength and tends toward a saturation point. Despite the proximal and negatively charged catalytic aspartates in the dimerized folded state, HIV protease does not exhibit halophilic properties at low or physiological ionic strength. Our results are plotted starting from 6 mmol/L ionic strength, calculated based on the buffer solution described in the corresponding experimental study[87].
Figure 2.3: Free energy of unfolding of HIV-PR as a function of ionic strength. A. Electrostatic stability profiles based on two independent apo wild type crystal structures of HIV-PR as a function of ionic strength. B. Unfolding free energy accounting for both electrostatic stability and the hydrophobic component. From the electrostatic point of
view, both of the structures show decreasing stability with increasing ionic strength with the curled structure (3PHV) being destabilized more than semiopen (1HHP). Error bars in B show the limits of the model using microscopic surface tension parameters at extremes (25 and 31 cal/ Å²mol).

Our results involving only the electrostatic stability (figure 2.3 A) of HIV-PR do not match with experimental values, which demonstrates an enhanced fold stability upon the addition of 1.0M NaCl [87]. It shows that the electrostatic component alone is insufficient in providing a complete description of the influence of salt on the stability of HIV-PR. Upon the inclusion of the salt dependent hydrophobic solvation model, we achieve better agreement with experimental results. The results achieved with the more complete salt-dependent model are shown in figure 2.3B.

Our results after adding the non-polar or hydrophobic component are in close agreement with the experimental data, once again demonstrating the importance of the salt effect through the hydrophobic mechanism. This can be seen directly by comparing results in figure 2.3 A and B for the salt effect of HIV-PR at 1.0 mol/L salt concentration. For clarity, it should be noted that this does not imply that the hydrophobic effect is primarily responsible for the stability and functioning of the enzyme, but simply that the influence exerted by NaCl on the stability of HIV protease is primarily described through Hofmeister effects.

Comparing figures 2.1 for CSP and figure 2.3 for HIV-PR, we can observe that difference between polar and non polar components of solvation free energy, in going
from 0.1 mol/L to 1.0 mol/L ionic strength, in the case of two HIV-PR structures (~14 kJ/mol) is almost 3.5 times as much than that for CSP family (~4 kJ/mol). This is expected as the surface area of HIV-PRs (~4200 Å\(^2\), ~3380 Å\(^2\) form monomer unfolding and ~900 Å\(^2\) upon dimer separation) and CSP family (~1200 Å\(^2\)) share a similar relationship. The surface tension function that mediates the relationship between the change in surface area and the non-polar solvation free energy is applied identically to both the HIV-PR and CSP systems.

4 The influence of salt on conformational equilibrium and activity of HIV-PR:

HIV-PR in its apo form is known to adopt a variety of functionally important conformations. NMR experiments have established the flexibility of the flap region, suggesting that closed, semi-open, and fully open conformations of the protease are in dynamic equilibrium, with the semi-open form being prevalent for the free protease [118]. Understanding this issue of flap mobility and the associated conformational equilibrium has profound implications on the mechanism and activity of the protease. Though the semi-open conformation is more populated, a curled conformation is also observed during the opening event in which the flap tips are curled back toward the protease burying additional hydrophobic side chains. Flap curling serves as trigger for flap opening and facilitates substrate binding [119, 120]. Our calculations indicate that salt effects, arising both from polar and non-polar contributions, preferentially stabilize the curled conformation (3PHV) relative to the semi-open conformation (1HHP). The enhanced activity of HIV protease observed at higher NaCl concentrations could be the
result of two effects. First, the enhanced stability of the folded conformation of the protease (a weakly stable enzyme under physiological conditions) at higher NaCl concentrations could lead to enhanced activity simply as the result of an increased population of functional enzyme. Second, the selective stabilization of the functionally important curled conformation could also contribute to the enhanced catalytic activity observed at higher NaCl concentrations.

5 Salting in and salting out behavior of salts in Hofmeister series:

After using our model successfully on CSP and then to explain the halophilic behavior shown by HIV-1 protease in presence of NaCl, we tested the same model to evaluate salting and salting out behavior of Hofmeister salts on CSP. The salting in or out effect exerted by salts in Hofmeister series also follow the same ranking for their surface tension increments per mole of salt in aqueous solution. The salts that salt out proteins effectively show larger surface tension increment and the salts that salt in proteins effectively show very low surface tension increment, as a function of salt concentration. Our model showed very good agreement with experimental results for different proteins (CSP) and enzyme (HIV-1 PR) which broadly pertains to the effect of NaCl on their fold stability. To accomplish this we used microscopic surface tension parameter as a function of NaCl concentration in the non-polar term. Similarly by using microscopic surface tension parameter for different salts in Hofmeister series, the same model can be made salt specific.
Figure 2.4: Salting in and salting out effect of salts A. Na$_2$SO$_4$ and B. NaSCN in Hofmeister series on fold stability of four structures in CSP family shown by solid lines. Dashed lines in A and B represent effect of NaCl on protein fold stability which are same in A as well as B and should be to be compared with same color solid lines.
Figure 2.4 clearly shows salting out effect of Na$_2$SO$_4$ and salting in effect of NaSCN compared to NaCl as detected by Hofmeister series [121]. As mentioned above in the text, surface tension increment as function of [NaCl] is 9.9 J/Å$^2$mol. The surface tension increment as a function of [Na$_2$SO$_4$] is 16.6 J/Å$^2$mol which is salting out salt and the surface tension increment as a function of [NaSCN] is 3.3 J/Å$^2$mol which is salting in salt of Hofmeister series [122]. Respective microscopic surface tension parameters are use to make the model salt specific shown in figure 2.4. Different salts alter surface tension of their aqueous solution at different rate. It mainly happens because of different strength of self hydration of different ions. The ions which have high charge density are strongly hydrated and show ion – water interaction favorable than water – water interactions. This is a behavior shown by salting out salts in Hofmeister series. Conversely the ions which have low charge density are weakly hydrated and show water – water interactions favorable and stronger than ion – water interaction. This scenario makes larger depletion in concentration of ions, in the layers close to interface, for strongly hydrated ions (salting out salts) than for weakly hydrated ions (salting in salts). According to Gibbs adsorption isotherm (equation 2.8) this depletion in number of ions near interface is the primary reason behind rise in surface tension. This is basic phenomenon behind existence of surface tension and increment in it at different rate for different salt in aqueous solution.

One of the proposed mechanisms, among few other predicted for origin of Hofmeister effect, is through change in surface tension at protein – water interface due to presence of salts. Salts that cause salting out, increase the microscopic surface tension at
this interface by greater magnitude in pre mole of salt than the salts that cause salting in. Strongly hydrated ions are excluded from the hydrophobic interface and are solvated in bulk solution phase favorably. This makes greater depletion in concentration of strongly hydrated ions (salts causing salting out) near protein – water interface and increase surface tension at the interface with very high magnitude per mole of salt concentration. The same fact also means that these salts will favor and stabilize more compact form of proteins (folded form) compared to exposed conformation (unfolded state). The weakly hydrated salt ions (salts causing salting in) favor the presence in close vicinity to the hydrophobic interface as they become excluded from bulk solution phase and increase concentration of ions near the interface. In turn these ions do not increase the microscopic surface tension to the extent of the rate at which salting out salt would do. Alternatively these ions are also shown to stabilize exposed conformation of polymers and peptides by direct binding to hydrophobic groups inducing polarization [121]. It means that salting in salts will favor and stabilize more exposed form compared to compact form of proteins. Along the lines of this mechanism, using different microscopic surface tension parameter for different salting in and salting out salts, one should be able to see difference in stability or unfolding free energy for salting out effect of salt compared to salting in effect of salt. This is the exact picture clearly seen from figure 2.4 where stability profiles for all the structures are higher for Na$_2$SO$_4$ and the same are lower for NaSCN compared to NaCl
Conclusions:

Here we have described a theoretical model to qualitatively and quantitatively evaluate the electrostatic effects and hydrophobic effects of salt on protein and enzyme fold stability. One unique aspect of this model is the inclusion of an ionic strength dependent microscopic surface tension used to address the influence of salt on the non-polar solvation energy. Utilizing this salt-dependent model of protein stability, we have demonstrated excellent agreement with experimental results determined for a substantial collection of 27 cold-shock protein variants. The model illustrates a balance whereby the influence of salt at low concentrations is exerted primarily through electrostatic screening, while at higher concentrations the effect is exerted primarily through non-polar solvation (often characterized as hydrophobic effects in aqueous solvents). Both of these physical mechanisms underlying the salt effect must be properly accounted for to properly describe the influence bulk ionic strength. Simultaneously our model also showed ability for predicting salt specific effects on proteins based on the results of salt effect of \( \text{Sa}_2\text{SO}_4 \) and \( \text{NaSCN} \).

The model was further applied to the HIV protease system in order to better understand the experimental observation of enhanced stability and activity in high (1M) concentrations of \( \text{NaCl} \). The results of the ionic-strength dependent stability calculations, applied to x-ray structures of the unliganded HIV protease, once again demonstrate excellent agreement with experimental measurements. In addition, calculations performed on two catalytically important conformations of the protease, the semi-open and curled conformations, indicate a preferential stabilization of the curled conformation.
at high NaCl concentrations. Prior work establishing the importance of this conformation in the initiation of catalytic activity suggests this preferential stabilization as a contributor to the enhanced activity observed at these conditions.
CHAPTER THREE

STUDY OF THE EFFECT OF SALT ON WATER STRUCTURE AROUND PROTEIN THROUGH EXPLICIT SOLVATION METHOD.

Abstract:

Although ion specific Hofmeister effects were identified and studied many years ago, the effect of different ions in close vicinity to protein surfaces remains a valuable area of study, and may provide important information regarding protein folding and evolution. One way of capturing Hofmeister effects on proteins is by quantifying how salts alter the association and dynamics of waters on protein surface. We examined this through molecular dynamics simulations of a model protein in a periodic box of TIP3 waters and salt ions ranging in concentration from 0.5 to 3.0 mol/L. Radial distribution functions g(r), self diffusion coefficient and dipole reorientational autocorrelation function of water in different shells from the protein surface as well as for water in salt solution was calculated. The coordination number was calculated for ions and water too. Based on these result we see an agreement between the rank for an ion in Hofmeister series correlates with its water structure making or breaking properties in protein, salt and water solution studied here. Conversely, the correlation times of water dipole reorientational autocorrelation function calculated in different shells, except the one which is between the bulk and hydration shell, show that the ions in Hofmeister series influences hydration water around protein similarly, based on their water structure
making and breaking properties. With our results we can articulate that effect of salt on the exchange rate of water could be important in mechanism of Hofmeister effects. The study serves as the basis for further investigations into the interactions of protein surfaces and solvent.
Introduction

Every one of us must have noticed a difference between the taste of “lite” salt, which is mixture of KCl and NaCl and ordinary table salt, which is essentially pure NaCl. In other words we have experienced specific ion effect in our daily life. Such effects are ubiquitous in chemical and biochemical processes involving salt solutions and have traditionally been attributed to influence of salt ions on the structure of water [123]. Few evidences in the past has provided facts that ion specific effects could be attributed to specific ionic interactions with surfaces and influence on hydrophobic interactions [123-125]. Since its discovery in late nineteenth century, Hofmeister effects have been examined in countless systems [126]. It includes salting proteins in and out of solutions, assembly and aggregation of hydrophobes like carbon nano tubes, water retention in macroscopic systems, molecular ordering of surfactants, and modulation of membrane protein conformations [121, 127]. The most commonly studied phenomenon has been effect of salts on protein stability and solubility [128]. It is now known that variety of process from enzyme activity to polymer and protein folding also display a Hofmeister Effects [129]. The salt ions studied for their Hofmeister effect are ranked in a series, which relates to the minimal concentration required to precipitate a given protein from aqueous solution at typically molar concentration. The ranking according to the ability of an ion to precipitate proteins yields following order of anions and cations [130].

<table>
<thead>
<tr>
<th>Most precipitating</th>
<th>Moderate</th>
<th>Least Precipitating</th>
</tr>
</thead>
<tbody>
<tr>
<td>PO₄³⁻ &gt; SO₄²⁻ &gt; HPO₄²⁻ &gt; Cl⁻ &gt; Br⁻ &gt; NO₃⁻ &gt; I⁻</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca²⁺ &gt; Mg²⁺ &gt; Li⁺ &gt; Na⁺ &gt; K⁺ &gt; NH₄⁺ &gt; Cs⁺</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Quantification of Hofmeister effect of a salt is usually obtained via linear regression of relative solubility or precipitation of a chosen solute generally proteins, peptides or small molecules, versus the salt concentration and slope of such correlation is termed the Setchenow constant [131, 132]. Although such constants provide useful measure of salting in and salting out behavior by indirect clues of salt induced changes in solvent’s bulk properties such as surface tension, viscosity and osmotic coefficient which can then be correlated with solubility of solute, they do not explain the mechanism by which salt actually modulate solute behavior in solvent. Changes in solution properties largely reflect changes in behavior in solvent, frequently water, therefore it is reasonable to imagine that salt influence the solubility of solute by altering the overall liquid water structure.

However, despite wide range of aqueous chemistries that follow the Hofmeister series, an underlying description of its mechanism at molecular level is still not complete. Two basic hypotheses are among the discussion in the field. One suggests that important source of ion specificity missed in classical double layer theory is ionic dispersion potential that acts between an ion and an interface. Ions have in general different polarizability (specific for each ion) than the surrounding water and hence experience a very specific a very different dispersion potential near interface [133]. At high concentration, where electrostatic potential become more and more screened, these ion dispersion dominate the interaction completely [69]. The other hypothesis involves various ion’s ability to make (called kosmotropes) or break (called chaotropes) water’s hydrogen bonding structure, based on ion’s interaction with water, the concept introduced
by Collins [134]. Structure makers show ion – water interaction favorable than water –
water interaction and structure breakers show the same to be weaker than water –water
interaction. However this idea has been called into question by recent experiments which
indicates limited influence of these ions on the structure and dynamics of bulk water i.e
beyond first hydration shell of salt ions [72, 135]. Analyzing effect of ions on structural
metrics of water, studies by Omta et. al. also concluded that ions classified as chaotropes
induced a more ordered solvent than those classified as kosmotropes and emphasize on
importance of strength of interaction between ions and water in their first hydration shell
[135, 136].

Never the less, the water molecules in first hydration shell of ions do not have
similar hydrogen bonding, in terms of the extent, geometry and dynamics that exist in
pure water, and it is beyond controversy. The problem lies in the water outside the
hydration shell and was tested by examining effect of ions on the reorientation dynamics
with ultrafast spectroscopy of water [135]. Many of such experimental determinations
require small water-to-ion ratios and the overlap of the hydration shells of the ions and
counter ions must be reckoned. In the place of kosmotropic/chaotropic paradigm, the
direct or indirect binding of ions to hydrophobic solutes or patches of proteins, and
associated disruption of water solvating these molecules, has emerged as possible
mechanism of Hofmeister effect. Kosmotropic ions were shown to polarize water
molecules hydrating the macromolecule, whereas the chaotropic ions directly interact
with macromolecule polarizing its atoms and stabilizing more solvent exposed form
[137].
Hydrophobic interactions are solvent induced. The characteristics of hydrophobic solvation that distinguish it from ‘common’ solvation, is the large negative change of unitary entropy at room temperature and large positive change in heat capacity can be explained by changes in water structure. The sign of partial molar constant pressure heat capacity \((\partial C_p/\partial P)_T\), indicates whether the solvated ion makes \([ (\partial C_p/\partial P)_T < 0]\) or breaks \([ (\partial C_p/\partial P)_T > 0]\) water structure [138]. An accepted explanation for large changes in these thermodynamic properties suggests that the water molecules around the hydrophobic solutes or on hydrophobic surface of a macromolecule arrange themselves in quasi crystalline structure or ‘iceberg’. These water molecules have entropy lower than bulk water molecules and approach the bulk entropy with rise in temperature and have positive partial molar constant pressure heat capacity. From thermodynamic perspective, effect of ion on salting in or out of solution was explained by Wyman in his theory of linked-functions [139]. According to his theory of linked functions a relationship exists between the change in chemical potential of a macromolecule upon addition of co-solute (salt ions) to the solution and excess binding/exclusion of that co-solute to/from macromolecule. This theory can also be derived from change in surface tension as a function of the excess amount of co-solute at the interface, a dependency give by Gibbs adsorption isotherm [140]. The preferential binding concept of co-solute has mainly been used to characterize changes in stability of native structures of protein with respect to unfolded states induced by addition of denaturing agents [141].

The Hofmeister series mainly pertains to the ions near the interface of interest which could be macromolecule or glass or solution of surfactant or monolayer of polymer
with solvent. In spite of partial resemblance of the Hofmeister series to the effect noted for ions at infinite dilution on the water structure, no well established correlation exist between protein stability and such water structural effect under the influence of salts. This conclusion of absence of definite correlation was based also on two state model of water that may but need not represent its actual structure nor the effects of ions on it [72, 142, 143]. Interfacial water molecules are already ordered non-isotropically by the surface at which they reside, so that the effect of ions on them cannot be directly compared with the effects in isotropic dilute aqueous solutions. The interfacial water structure are affected by presence of salts and degree of ordering was shown to follow Hofmeister series as directly observed spectroscopically for octadecylamine monolayer [121]. The same should be true for macromolecules like proteins in aqueous solution, with the surrounding water molecules having already some structure forced on them by presence of the macromolecule with hydrophobic and hydrophilic patches on its surface, so that the ion effects described by Hofmeister series should differ from those in homogenous isotropic dilute solutions. It would be valuable and interesting to examine effect of high concentration of salt on the waters at protein – water interface (hydration water) and may hold possible key to the mechanism of modulation of water on surface as done in this study.

Salting out of small hydrophobic molecules like methane and neopentane through their hydrophobic self association correlates well with effect of salt on bulk water structure. This was concluded by Elcock by calculating relative number of hydrogen bonds per water molecule from molecular dynamics simulations salt water and small
hydrophobic organic molecule solutions [144]. The system of such small organic molecules can closely represent unfolded state of proteins where peptide chains are more exposed to solvent compared to folded or native state. In the work by Koga et al, the relative effects of Hofmeister anions on salting out 1-propanol are well explained by their water structure affecting abilities and are shown to have no relation to surface effects and Hofmeister series [145]. It appears from such results that for small hydrophobic solutes or other hydrophobic solutes on similar length scale, water around them is not as strongly structured as near macromolecules like polymer chains, proteins or enzymes which has larger surface area. The water structure affecting properties of ions can then be examined more clearly around macromolecules by studying these molecules immersed in saline solution.

A large number of studies have used molecular dynamics simulations and experimental techniques to explore behavior of ions and its impact on solubility of solute in aqueous solution to resolve questions about the origin of Hofmeister effects [130, 144, 145]. The majority of these studies have focused on pure salt solutions or small hydrophobic/hydrophilic solute molecules in aqueous medium in the presence of salt ions. They have been primarily aimed at hydration thermodynamics of constituent ions and solutes and correctly describing structures of their ion’s hydration shell. Other studies have focused on behavior of ions at vapor-water interface and have shown that the localization of ions at the interface depends strongly on nature of ions. Such studies also give insights into surface tension effects which often correlate with Hofmeister effects [71, 125, 146].
These studies of pure aqueous salts and small solute molecules or hydrophobic objects provide important groundwork for studying more complex issue of how proteins behave when immersed in salt solution. One way of investigating such behavior is to examine the effect of salt ions on water in close vicinity of protein surface and the changes in protein conformation. Water near the biomolecular surface behaves differently than bulk water. The entropic penalty enforced on water near protein surface arises from orientational/rotational as well as translation restriction on water. The effects of ions on biological water cannot be directly compared with the effects on isotropic dilute ionic solution i.e bulk water. Indeed the effect of ions on biological water is important area for studying Hofmeister effects on the stability and solubility of proteins and enzymes. In this work, we study the behavior of ions at protein surface - water interface and their influence on structure water in the immediate vicinity of protein surface.

In this study we examine the structure of water and influence of salt on it near protein surface. In particular, in this work we focus on effect of salt on hydration water and compare the same effect on the bulk water. We attempt to learn how different the effect on hydration water is from that on the bulk; thereby possibly shine light on mechanism behind Hofmeister effects in case of proteins. We begin with investigating the influence of salt on structural properties of water from simple salt ions in water to more complex system of protein and salt in water to examine protein solvation. As a comparison, we also study neat water and protein in water system. First we test our simulation protocol by comparing structural elements calculated for neat water with that from experimental neutron diffraction data and other theoretical calculations. Chloride
salts of magnesium (MgCl$_2$), sodium (NaCl) and cesium (CsCl) were chosen for this study. Each salt has been simulated in explicit water solution using TIP3P water model. Properties like radial distribution functions of water – water as well as water – cation interaction yield the information about long and short range water structure. The effect of salt on water’s translation and rotational motion was studied through calculating self diffusion coefficient via mean square deviation and water’s reorientational autocorrelation function decay time. Simultaneously we make use of properties like hydration number of cation and cation-water and water-water hydrogen bonding lifetime and number of hydrogen bonds to access the effect of salt on hydrogen bonding pattern of water. Then we focus on more complicated phenomenon of influence of salt on hydration water. We selected two proteins from the cold shock family that differs in net charge. We study the solution of protein-water as well as protein-salt-water.
Materials and methods:

1 Simulation systems and protocol:

To examine the effect of salt on protein and water structure around it, molecular dynamics simulations have been performed for four different systems using CHARMM27 all hydrogen force field [147]. It includes neat water, salt ions – water, protein – water, and protein – salt ions – water system. Salts studied in this work include MgCl₂, NaCl and CsCl. TIP3P water model was used as water solvent [148]. All the salt ions were modeled using parameters available within CHARMM27 force field package. Parameters for Mg²⁺ and Cs⁺ within CHARMM27 are untested for usage in large number of different simulations. This study also confirms their usage for calculating properties of interest shown here. Simulations for neat water and salt ions - water system were performed using cubic box of edge length 30Å to which periodic boundary conditions were applied. Salt ions – water systems was studied for 1, 2 and 3 mol/L of salt concentration. For the protein – water and protein – salt ions – water systems cubic box with edge length of 54Å was used. Crystal structures of folded proteins from cold shock protein family, derived from the mesophilic bacterium *Bacillus subtilis* (Bs-CspB) PDB:1CSP [89] and the thermophilic bacterium *Bacillus caldolyticus* (Bc-Csp) PDB:1C9O [90] are utilized in this study. Water molecules with oxygen atom within 2.6Å of any atom of protein were deleted after putting protein at the center of water box. The protein – salt ions – water systems were studied at 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 mol/L of salt concentrations. 0.1 mol/L NaCl concentration was maintained for protein – water system
to simulate standard biological salt environment. The cold shock proteins studied here carries \(-6\) and \(-2\) net charge for Bs-Cspb and BcCsp respectively. To neutralize the system, appropriate number of respective extra counter cations was added. Migration of the solute protein outside of the primary solvent box was discouraged during the molecular dynamics simulation by weak \((0.5 \text{ kcal mol}^{-1})\) center of mass translational restraint using MMFP module of CHARMM [149] on all alpha carbon atoms of protein.

Water by itself is \(~55\text{M}\). Hence replacing one water molecule by a salt molecule (pair of ions), per 55 water molecules yield \(~1\text{M}\) salt solution in water. Following such method salt ions were added by replacing randomly selected water molecules in pre-generated water box (or water protein box). SHAKE algorithm was used to constrain all bonds involving hydrogen [150]. Systems were first energy minimized with steepest descent minimization for 100 steps and ABNR minimization for 400 steps.

A 12Å cutoff was used for particle-particle interactions, and non-bond list was constructed by using 16Å cutoff. The long range electrostatics was treated with particle mesh Ewald (PME) method with 10Å cutoff. B-spline order of 4 and real space Gaussian width \((\kappa)\) of 0.34 Å\(^{-1}\) was used with PME [151]. The system was slowly heated from 100 to 300 K over the course of 200ps and then was equilibrated for another 700ps followed by 1ns of production run. As indicated in the text following results and discussion, small and non overlapping error bars on number of properties plotted confirms that simulation to be adequate for properties of interest in this work. All the simulations were run at constant pressure of 1atm \((101,325\text{Pa})\) and temperature of 300K using 2fs time step. The temperature of system was kept at 300K on average with Nose-Hoover thermostat [152,
and constant pressure was maintained by the Langevin piston method [154].

2 Computing molecular association:

To access the effect of dissolved salts on liquid water structure in the simulations, different measures were used. The first was familiar water oxygen – water oxygen radial distribution function (RDF), which describes the spherically averaged distribution of oxygen – oxygen separation distances between pairs of water molecules sampled from the simulation trajectory. Calculating such RDFs for water oxygen – water hydrogen would describe hydrogen bonding pattern and that for cation – water oxygen served as an indicator of cation solvation shell strength.

Association of water around proteins was examined at various salt concentrations to account the influence of salt on it. Hydration number (number of water molecules around species of interest) for cations were defined as number of water oxygens within certain radius as function of distance from center of the solute. The lifetime of water – protein and water – cation and water – water hydrogen bonding too, was calculated using 2.4Å cutoff for water oxygen – water hydrogen and 3.5Å for water – protein and cation – protein interactions.

3 Effect of ions on translation and rotational motion of water:

To measure influence of salt on hydration water (water very close to protein surface) and bulk water, self diffusion coefficient and P1 reorientational correlation times were calculated for water in different shell from protein surface. Self diffusion coefficient
were calculated from mean square displacement (MSD) of all water oxygen atoms using the Stokes-Einstein relation

\[
\lim_{t \to \infty} \langle |r(t' + t) - r(t')|^2 \rangle = 6Dt = \frac{k_BT}{\pi R_h \eta} \tag{3.1}
\]

where \( r(t) \) is the position of water oxygen at time \( t \), \( R_h \) is hydrodynamic radius of solvent particle, \( \eta \) is viscosity, \( D \) is self diffusion coefficient, and the brackets denote averaging over all water molecules at time origins \( t' \). The self diffusion coefficient was estimated from the slope of the linear part at long time of mean square displacement versus time plot. Along with water molecules, self diffusion coefficients of cation were measured too.
Results and discussion:

1 Radial distribution functions for water structure:

The radial distribution functions (RDF) are commonly used when structure of liquids or gases are studied. These intermolecular partial pair correlation functions were calculated for water – water (gr_OO, gr_OH and gr_HH) as well as water oxygen – cation (gr_cation – water (O)) interactions. We evaluate the difference in water structure around ions by using radial distribution functions and coordination numbers for ions.
Figure 3.1: Radial distribution functions (RDFs) calculated for TIP3P water oxygen-oxygen gr_OO, oxygen – hydrogen gr_OH and hydrogen – hydrogen gr_HH interactions in different salt solutions at different concentrations. The last column shows RDFs for neat TIP3P water simulation (purple line) and that from the experimental neutron diffraction data (orange line) from Soper et al [155].
The radial distribution functions for TIP3P water in 1, 2 and 3 mol/L salt solutions of MgCl₂, NaCl and CsCl are calculated and compared with those for neat TIP3P water and neutron scattering experimental data in figure 3.1. Heights of first maxima for gr_OO are summarized in table 3.1. Our calculated radial distribution
functions for neat TIP3P are in good agreement with previously reported results in literature. The first peak positions for neat TIP3P water calculated in this study is at 2.8Å which matches well with previously reported value in literature at 2.77Å [156]. Our peak position for gr_OO occurs at shorter distance when compared with experimental neutron data value which is 2.88 [155] but it is closest to previously reported theoretical position. The height of the peak is slightly lower compared to experimental result, which was also seen in other theoretical studies. This indicates the current TIP3P water model produces its first hydration shell weaker then observed in experiments. Also structure beyond first peak is missing in a way that height of second peak too low and position is shifted to shorter distance. This indicates that second hydration shell of water, which should be related to self diffusion coefficient, is lesser than experimental finding. To have too little structure in gr_OO RDFs and lack of structure beyond the first peak, the difference between calculation and experimental could be attributed to currently available accuracy in determining the site-site pair correlation functions of water and to the well-documented problem for the TIP3P model [156-158]. Fluctuating charge polarizable model and force field is shown to overcome some but not all discrepancy in this issue and have indicated to be possibly overcome with optimized LJ 12-6 potential [159] but their extra computational cost and compatibility issues with all types of co-solutes salt ions and protein make their usage limited. In addition current work also show level of precision and accuracy that can be attended in salt effect related studies with commonly used water model and simulation protocol.
For gr_OH, the peak height occurs too short as compared to the experiment. Comparing gr_OO peaks among three salts, as salt concentration increased, it is increased by presence of Mg$^{2+}$ while it is reduced by presence of Na$^+$ and Cs$^+$. This is result of disruption of local hydrogen bond network by presence of Na$^+$ and Cs$^+$ while increased in hydrogen bond network by presence of Mg$^{2+}$ ions. This is expected as Mg$^{2+}$ is structure inducing kosmotrope and significantly reorders local hydrogen bonding while Cs$^+$ is structure breaking chaotrope and induces disorder in the hydrogen bonding network of water. Alternatively, gr_HH increases for Cs$^+$ than other two cations. This is because Cs$^+$ being poorly hydrated ion, forces water molecules to form cage like structure with cavity for Cs$^+$. Comparing gr_OO and gr_HH, the later is highest for CsCl solution while former is lowest for the same salt. This may be due to hard, short-range repulsion implicated with the LJ 12-6 potential. The weak magnitude of LJ interaction between hydrogen – hydrogen, compared to other heavy atoms makes gr_HH immune from water structure making effects of Mg$^{2+}$. When water favor thermodynamically preferred (highly favorable enthalpy through maximized hydrogen bonding pattern and unfavorable entropy) cage like structure around Cs$^+$, unlike around Mg2+ where it is octahedrally coordinated in first shell, the enforcement in gr_HH is seen.
Figure 3.2: Radial distribution functions (RDFs) calculated for cation – TIP3P water oxygen interactions $g(r)_{\text{cation-water(O)}}$ (top row), hydration number of cations (middle row) and reorientational autocorrelation function of water dipole (bottom row) in different salt solutions at different concentrations. Error bars of standard error are shown by vertical lines capped by horizontal at each data point. Legend shown for $g(r)_{\text{cation-water(O)}}$ 2mol/L is same for all plots.
2 Cation – water interactions:

Dissolved ions in water are surrounded by water molecules with variable degree of order that depends on the nature of the individual ion. These ions hold their hydration shell with strength that too depends on nature of ion. The interactions of ions with waters in their first solvation shell are charge-dipole interaction. To compare interaction of cations with water, we evaluate difference between the water structure and the cation-water radial distribution functions. Seen in figure 3.2 first row is gr_cation – water oxygen. The height of first peak diminishes with decrease in cation’s charge density from $\text{Mg}^{2+}$ to $\text{Cs}^+$ and implies a weakening of cation-water interaction. Also the position of first maxima shifts towards longer distance as cation size increases and charge density decreases from $\text{Mg}^{2+}$ to $\text{Cs}^+$. This represents the first solvation shell strength of these ions which is dominated by charge to size ration of the ion. We can see that there is profound second peak for $\text{Mg}^{2+}$ and that for $\text{Na}^+$ which is little further as compared to $\text{Mg}^{2+}$. This confirms that $\text{Mg}^{2+}$ holds second hydration layer strongly and with shorter radius of second hydration layer compared to $\text{Na}^+$. $\text{Cs}^+$ lack in the second peak and indicates absence of second hydration shell for it. The same plot also yields the information about effective radius of these cations with their hydration shells which is closely related with diffusion coefficient of cations. Our results for cation –water oxygen RDF matches well with previously reported experimental as well as separate theoretical RDFs for $\text{Cs}^+$ $\text{Na}^+$ and $\text{Mg}^{2+}$ respectively with less than 4% deviation [160, 161].

The 2$^{nd}$ row in figure 3.2 shows hydration number or number of water molecules as function of distance from cation. The strength of hydration shells shown from top row
is consistent with the hydration number of cations seen in middle row. It represents data about effectively how much of volume is occupied by presence of cation and his hydration shell and how many water molecules surrounds that volume, in other words hydrodynamic radius of cation. Cs\(^+\) being largest of cation with large ionic radius (1.81Å) occupies a large volume but with very low charge density does not hold hydration shell to the extent Mg\(^{2+}\) holds. On the other hand Mg\(^{2+}\) and Na\(^+\) (radius 0.72 Å and 1.16 Å resp.) occupy smaller volumes with higher charge densities but Mg\(^{2+}\) can hold almost two hydration shells very strongly and effectively moves as a larger particle.

![Figure 3.3](image.png)

**Figure 3.3:** Dipole reorientational autocorrelation function decay time of water in salt solution as function of salt concentration.

3 Rotational motion of water:

The 3\(^{rd}\) row in figure 3.2 in combination with figure 3.3 shows reorientational
autocorrelation function decay time of water dipole. The rotational correlation of water has been studied for many years by theory as well as experiment [158, 162, 163]. The water dipole reorientational autocorrelation function describes how long (in terms of time) a water molecule or a group of water molecules can preserve their initial dipole orientation. The dipole autocorrelation function of water in salt solution is compared with that for neat water. As seen in 3rd row of figure 3.2, the neat water correlation function decays faster than water in all the salt solutions showing shorter decay time. It suggests that the water in pure form is rotating faster and is free to do so than water in presence of salt ions. It follows the trend shown above by strength of hydration shell and in turn the rank of a cation in Hofmeister series and ability of an ion to polarize water. Tighter and bigger the hydration shell for Mg\(^{2+}\) than other two cations shows more water is confined around the cation and is less free to rotate yielding slower rotational dynamics and slower decay in reorientational autocorrelation. Na\(^+\) and Cs\(^+\) still shows slower reorientational dynamics of water compared to neat water because, though weak, but polarization effects, enforcing little more structure on water. This effect of Na\(^+\) and Cs\(^+\) on polarization of water is less compared to Mg\(^{2+}\). For Na\(^+\) and Cs\(^+\) there is very little difference of decay time at high salt concentration (3.0mol/L) compared to neat water. Our calculated reorientational correlation times of neat TIP3P which is 2.55 ps matches with reasonable accuracy as reported before for TIP3P. The neutron diffraction experimental values have been shown to be converged to about 2.0ps in literature [163]. Thus this picture confirms the effect of salt on rotational motion of bulk water. It is clear that water is greatly hindered in rotation motion by presence of highly polarizing salt ion Mg\(^{2+}\). In the protein
hydration when water is restricted in its rotational and translation motion on protein surface, it is important to compare the effect of salt on such hydration water to that on bulk. This is shown in detail further in this study.

Figure 3.1, 3.2 and 3.3 shows overall impact of salt on the structure of water. With these results cations can be now be compared among themselves for their properties of water structure making and breaking. Results suggests that hydrogen bonding between water molecules is made stronger by presence of \( \text{Mg}^{2+} \) compared to \( \text{Cs}^+ \) and \( \text{Na}^+ \). We demonstrate that using 1ns simulation, we could reach the same results for cation water interaction previously reported in literature from different theoretical as well as experimental studies. By validating our protocol by comparing to other theoretical and experimental results for cation – water interactions, referred at appropriate places above, we have shown modulation in water’s translation and rotational dynamics upon addition of salt. Many of previous theoretical calculations concerning cation – water interaction were carried with different set of parameters, force field and water model. Comparison of different properties of cation – water interactions calculated here, with such previous theoretical and experimental results, also confirms the compatibility of force field, water model and set of parameters used for salt ions in this study. These new results of effect of salt on water – water interactions in the solution of salt concentrations ranging from 1 to 3 mol/L, will serve reference for comparison between aqueous salt solutions and aqueous salt and protein solution. Further we apply similar simulation protocol to a system of protein and salt ions solvated in a box of water. We simulated two proteins from cold shock protein family derived from mesophilic bacterium \textit{Bacillus subtilis} (1CSP net
charge -6) and the thermophilic bacterium *Bacillus caldolyticus* (1C9O net charge -2). We simulated three salts used above to see the effect of salts on structure of water near protein surface.

Figure 3.4: Self diffusion coefficient of water in different systems. The colors of lines show different salt solutions as denoted by legend. The solid lines denote 1CSP-salt-water system and dotted lines denote 1C9O-salt-water system. Dash-dot lines denote different salt-water solutions. The orange triangle and circle shows data for protein-water system where 0.1mol/L of NaCl was maintained to match mesophilic environment. Error
bars of standard error are shown by vertical lines capped by horizontal at each data point.

4 Self diffusion coefficient of water and cation:

Self diffusion coefficient of water was evaluated using linear portion of the MSD, mean square deviation of the position of oxygen of the water molecule from starting reference position vs time plot shown in figure 3.4. It was calculated for water in all different systems which include neat water, salt-water solution, protein-water system and protein-salt-water solution using Einstein’s relation where the slope of MSD vs time (at long time) is 6D. While reorientational correlation of water yields information about effect salt on waters rotational motion, self diffusion coefficient yields information about the same factors on water’s translational motion. The self diffusion coefficient were reported previously in literature for TIP3P water model between 0.52 to 0.7 Å²/ps [158] using similar protocol as used in this study. Both these values are again off from experimental value of 0.23 Å²/ps [164] which is documented and explained in various TIP3P water model literature. Our calculated value for D of neat water is in good agreement with these previously reported values for TIP3P water model.

First we note that water lose its mobility by presence of a protein as compared to neat water. This is due to some of the waters near to protein are influenced by protein surface and is restricted in its motion. This effect is exerted mostly on hydration water but also extends upto 7 to 10 Å from protein’s surface. Also as seen figure 3.3, water dynamics is slowed by presence of all the salts and this effect is increased with increase in charge density of the cation. In addition water in presence of protein and salt (protein-
salt-water system) is less mobile than water in salt solution (salt-water system). The self diffusion coefficients decrease as salt concentration increases. This is due to increase in viscosity of the solution and follow a linear trend. In all the cases diffusive ability of water is hampered by presence of vicinity of protein though the difference is small. A bigger difference is occurs in waters mobility is due to presence of different salts. This shows that a small fraction of water molecules are bound to protein surface and behaves as hydration water. Overall effect of salt on diffusivity of hydration water is not much different than on bulk water. In general, though hydration layer of water show 10-15% increase in viscosity compared to bulk, only 5-30% of hydration waters exhibit reduced motion. This is confirmed through our result of self diffusion coefficient of water in vicinity of protein as well as in the bulk. It also throws light on the fact that water translation motions are not central to the Hofmeister mechanism.
Figure 3.5: Self diffusion coefficient of cation in different systems. The colors of lines show different salt solutions as denoted by legend. The solid lines denote 1CSP-salt-water system, dotted lines denote 1C9O-salt-water system and dash-dot lines denote different salt-water solutions. The orange triangle and circle shows data for protein-water system where only 0.1mol/L of NaCl was maintained to match proteins natural environment. Error bars of standard error are shown by vertical lines capped by horizontal at each data point.

Further we also calculate diffusion coefficient of cations in various simulation systems. These results are presented in figure 3.5. As expected, the mobility of ions
accessed through their self diffusion coefficients follows their abilities to bind solvation shell around them. Cs\(^+\) could be more dynamic than Na\(^+\) and Mg\(^{2+}\) as it is weakly hydrated than other two cations. Simultaneously cation’s dynamics is hampered in vicinity of protein. With figure 3.4 and 3.5 where water too, is seen to be less mobile in the vicinity of protein, effect of increased viscosity of medium is seen on mobility if cation additionally. The hydrodynamic radius is the effective radius of an ion and water in hydration shell as it migrates through the solution collectively. In other words this is a radius of solvated ion and it depends on the forces between ion and solvent. If these forces are sufficiently strong then ion retains larger solvation shell around it and moves effectively as a body with larger radius than crystal or gas phase radius of that salt ion. It follows that a cation with greater charge density will retain solvation shell more tightly and there can be multiple solvation shells. Ions like these always have largest solvation shell than the one with lower charge density. One of the hypotheses behind Hofmeister series is direct binding of cation causing salting in behavior. Large cations with low charge density are less hydrated and essentially are excluded from solution phase towards hydrophobic boundary. These cations also show better mobility than cations with high charge density. Seen from figure 3.4, surrounding of a protein has highest impact on Cs\(^+\) (comparing green dash-dot line which represents salt – water solution against solid and dot lines, both representing protein – salt – water solution) but it still holds highest mobility among three salts. Furthermore in conjunction with figure 3.4 and 3.5, the results of self diffusion coefficient of water and cation it can now be said that the dynamics of water molecules or cations does follow the trend in Hofmeister series but
possibly it is not the governing factor in mechanism.

Small ions like Mg\(^{2+}\) having high charge are strongly hydrated, with small negative entropies of hydration, creating local order and high local density. In contrast, large ions like Cs\(^+\) having single charge, with more positive entropies of hydration act like hydrophobic molecules, binding to surfaces dependent not only on charge but also on van der Walls forces. Additionally, they are pushed towards protein – water hydrophobic interface by strong water – water interactions and certainly induce a change in the surface hydration and interfacial aqueous clustering. Such large ions possess low surface charge density like Cs\(^+\) and are able to sit comfortably within dodecahedral water cage and produce lowest apparent density of the solution water. Smaller ions cause partial collapse of such cage structure through puckering as these ions tightly hold strongly to the first shell of hydrating water molecules and hence there is less localized water molecule mobility and higher apparent density for the solution water. This is evident from figure 3.2 and 3.3. A similar evidence was reported in case of artificially neutralized ion and charge ion [165].
Figure 3.6: Average number of hydrogen bonds and average life time of hydrogen bonds per water molecule, for water in different system studied. The colors of line denote presence of different salt in solution. Solid lines are for ICSP-salt-water and dotted lines are 1C9O-salt-water. Error bars of standard error are shown by vertical lines capped by horizontal at each data point.

5 Hydrogen bonding between water molecules:

Although it is generally accepted that the water is highly structured liquid, there is a little agreement on how to access the structure quantitatively and how to measure it experimentally. The experimental methods commonly employed to ascertain the structure of liquid, namely X-ray and neutron diffraction yield the structural factors and indirectly after Fourier transformation the total pair correlation function. Theoretical simulation methods can too calculate such properties reasonably matching well with experiments. It is also generally accepted that outstanding cause of structuredness of water is its extended hydrogen bonded network. A quantitative measure of this extent, namely the average number of hydrogen bonds and average lifetime of these hydrogen bonds per water molecules present in liquid, is therefore a very useful measure for the structuredness of water, in particular also as an answer to the question of how ions affect this structure. It is recognized that the regular tetrahedral hydrogen bonded structure of ice, which has four hydrogen bonds per water molecule, is broken down on melting. The average coordination number of water molecule in cold liquid water is somewhat more than four so there must be molecules not fully hydrogen bonded to other in addition to those that
are so bonded.

Shown in figure 3.6 are average number and average lifetime of hydrogen bonds per water molecules. The data is truncated at first decimal as default option of CHARMM. It can be quickly seen that lifetime of hydrogen bond per water molecule is increased due to presence of solute protein and salt compared to neat water while average number is decreased drastically for water by presence of Mg$^{2+}$ cations and protein. This is due to polarization and structuredness induced on hydration water. Furthermore the strong polarization effect of salt ions is seen from both of the graphs. As salt concentration increases, strongly hydrated cations like Mg$^{2+}$ makes less and less free water molecules in the solution there by reducing the number of hydrogen bonds per water and increasing average lifetime of hydrogen bonds per water molecule. This structure making effect of Mg$^{2+}$ is clearly seen from these plots and happens to the maximum at 1.0 mol/L of concentration. Cs$^+$, in contrast, is able to alter the structure of water in terms of hydrogen bonding pattern only after concentration 1.5 to 2.0 mol/L.

In the efforts of exploring mechanism of Hofmeister effects, although some results indicate that salt ions destroy natural hydrogen bonding pattern of water similar to the effect of change in pressure and temperature [166], some methodology suggests that ion cause negligible change in water’s bulk structure [135, 137]. These differences may be attributed to ionic concentrations used, the sensitivity of method to the scale of potential structural changes in the bulk water, the difficulty of separating coexisting but opposite effects of chaotrope and kosmotrope ions, precise meaning of bulk water and the importance of presence of surface in stabilizing effects. The last reason is particularly
important in case of effect of ions on protein stabilization. The effect of salt at lower concentrations may be smothered by relatively large amount of unaffected bulk water present whereas at high concentrations there may be insufficient water to properly show any specific effects. In light of explaining mechanism of Hofmeister effects, the present study examined effect of salt ions on mobility and dynamics of water. Particularly we were interested in translational and rotational motion of water. To overcome some of the above mentioned possible difficulties, we tested effect of salt on water in various systems. We included simple salt-water solutions ranging from 1 to 3 mol/L concentrations and compared to neat water for studying effect of salt on bulk water. Further we also studied protein solvated in water and compared effect of salt on water around protein by studying protein-salt-water system.
Figure 3.7: Dipole reorientational autocorrelation function decay time of water in protein-salt-water solution. Different colors of lines denote reorientational autocorrelation time calculated for water in different layers from protein surface. Layers are differentiated as denoted in legend. Two proteins are studied for this result. Each plot represents correlation function time of water in different salt solution around protein.
6 Water rotation dynamics in different layers from protein surface:

In the end to access the effect of added salt water in different layers from protein surface, we calculated properties of water namely dipole reorientational autocorrelation function times in different layers thickness from protein surface. Biological water or also called hydration water that resides close to as far as 10Å from the protein surface, behaves markedly different than bulk water [167]. This water mainly serves purpose of hydrating protein by directly interacting with residues on the surface through bridging interaction or act as legend. Different categories of water behavior can be characterized depending on their residence time on protein surface. One type is that strongly bound to protein, which can be identified crystallographically and play an important role in stabilizing the native structure. The residence time of these water molecules are usually in the range of $10^{-9}$ to $10^{-3}$ s. the other type of water is more dynamic with shorter residence time and is of interfacial nature. The third type of water is in the vicinity of protein surface but not directly interaction with the protein and behave mostly like a bulk. Rest all the water that do not interact with protein and free to ‘wonder’ which are considered bulk water. Hofmeister effect have been shown in context of behavior of water in vicinity of salt ions (pertaining to ion hydration) or in vicinity of small solute molecules that are dissolved in water along with salt. Our main interest here is to elucidate the effect of salt on water in close proximity of protein is a different and new approach to the interpretation of Hofmeister effects that we employ through this study.

Figure 3.7 shows that as we go away from protein surface the the dipole reorientational autocorrelation function decay becomes more andmore quicker. Authors
note that a more rigorous way would be to take contribution of all waters rather than to account for only water molecules that did not diffuse into other layers, and such method of calculating reorientational autocorrelation functions in layers could result in more precise data of current analysis.

The layer of water in closest proximity 0 to 3 Å from protein surface shows very high correlation times in which water is bound to surface. These water molecules show decay time of their reorientational autocorrelation function 1ns as shown in figure 3.7. Considering simulation length used here which is 1ns, we report these water will have rotational correlation function life time $\geq$ 1ns. The next layer 3 to 6 Å away from protein surface, show dipole reorientational decay times shorter which mean these water molecules are rotating faster than the closeese layer. The layer most away from protein surface, 9 to 12 Å, shows reorientational autocorrelation function time similat to that of bulk. Comparing this picture accross all three graphs, which represents the effect of different salts, there is very little difference in effect of different salts on water dipole orientation, except the layer 6 to 9 Å where it fluctures by largest amount from salt to salt. Possible reasone behind this perticular layer showing highest flucturaions in decay times as this is the layer where water is on the brink of being hydration water or bulk water. The exchange rate for water between bulk and hydration water is highest in this layer. Calculating properties of water in layer of restricted volume from peotein surface is lsee straight forward or not well defined than interpreting the same property for bulk.

When analyzing property of water in a shell, water molecules contribute to the property of interest only when they are in that shell. The effect fo salt on water in this shell is
different than bulk. In this shell water dipole reorientational autocorrelation function times attends lowest value at smallest concentration for Mg$^{2+}$ (1.5 mol/L) compared to Na$^+$ (2.0 – 2.5 mol/L) and Cs$^+$ (2.0 mol/L). Such phenomenon is seen in the layer 6 to 9 Å away from protein. With these results we can imply that the exchange rate of water between hydration layer and bulk water is been affected the most by presence of salt and that it might be a key behind mechanism rather than pathway of altering structure of water alone.
Conclusion:

The results shown here present important new insights to the observations we made on salt – water, water – water interactions through number of protein-salt-water systems using solvation dynamics. We have examined effect of cations, from the most and least precipitating positions in Hofmeister series, on solvent TIP3P bulk water structure and water structure around protein surface. We calculated radial distribution function, water’s translational and rotational dynamics via self diffusion coefficient and dipole reorientational autocorrelation function and also the effect of ions on rotational motion of water in different layers from protein surface. The results have shown excellent agreement between changes in water structure upon affected by presence of the cation and cation’s rank in Hofmeister series. Structure making effect of divalent ion Mg$^{2+}$ is seen very effectively than strong structure breaking effect of Cs$^+$ is seen though it follows the rank order. Though Cs$^+$ showed results similar to Na$^+$ ions in modulating many of the solvent properties, there is measurable difference in the effect of these two ions on water. The salts showed their water structure making or breaking effect to the maximum extend on the water in the interfacial layer between hydration water and bulk water. Overall the present analysis suggests that at least a portion of physical basis for Hofmeister effects of salts on proteins is through modification of solvation environment and mainly on the layer of solvent between the bulk and hydration layer, as seen from effect of ions on water structure around proteins.

Given that efforts to understand the Hofmeister effects of salts often focus exclusively on the interaction/association between salt and the solutes like small organic
molecules or hydrophobic objects, it is important to note that we have been able to show here that qualitative prediction of Hofmeister effects can be made from different solvent properties calculated here from simulations of protein-salt-water system. The key result confirms that the change in water structure due to addition of salt may be fundamental to the Hofmeister effects than preferential binding. Also the effect of salt on different layers of water from protein surface is seen to be exerted to similar extent, except one layer where water is being continually exchanged between hydration layer and bulk. The most exaggerated effect of salt was seen on water molecules in the layer that is on the brink of being in the hydration layer and bulk. These results offer no direct evidence of specific site binding protein-ion interactions, but if specific interactions do occur, it is in conjunction with changes in water solvation properties. At the same time we comment that such events can be monitored with longer timescale molecular dynamics simulation of same kind.

Although our phenomenological study captures some important aspects of Hofmeister effect, the actual water dynamics and salt effect on protein stability is likely to be more complex. For example at some binding site, the bound to free transition rate of water as well as cations may vary with time depending on specific configuration of the local environment of the hydrogen bond network and concentration of salt. These events may have lifetime much longer than average calculated in this study and may need much longer trajectories to study such events. In these transition rates of these exchanges it self can become a dynamic property.
CHAPTER FOUR

COMPARISON BETWEEN IMPLICIT AND EXPLICIT SOLVATION METHODS:
EFFECT OF VISCOSITY ON SAMPLING PEPTIDE DYNAMICS

Abstract:

Computer simulations are used to explore the basis of structure and function of biomolecules. Presently available simulation methods involve explicit as well as implicit representations of solvent, each with their own pros and cons. The most noticeable difference is viscosity and stochastic collisions of water molecules in explicit solvation method which completely absent in implicit solvation. We carried a comparative study of the GBSW implicit solvent model with explicit solvation using TIP3P water on model with solute peptides poly-Proline type II helix and poly-Alanine alpha helix, and Cold shock protein A, for solvation model’s efficiency as well as accuracy in sampling trajectory. Normal model analysis was used to confirm long timescale, large magnitude and low frequency motions involved these solute structures. Our results showed a significant difference in efficiency of sampling conformational phase space to reach the same equilibrium state and it depends on solvation model used as well as on the cross section of type of secondary structure or protein under study. To reach the same equilibrium, GBSW proved faster by factor of 10 in sampling trajectory than TIP3P explicit solvation. Simultaneously solvent also has an effect of its viscosity on correlated motion of solute protein or peptide. Correlated motions of solute are damped by solvent
with high viscosity and similar to protein flexibility. Most damping effect of solute correlated motion by viscosity of explicit solvent occurs for solute with smaller cross section. An additional modified explicit solvent, based on the TIP3P model and higher only in viscosity by factor of 10 than TIP3P was studied too. Our results confirmed that the enhanced sampling efficiency and speed of GBSW was originating from an inherent lack of viscosity and stochastic collision of solvent molecules with solute compared to explicitly solvated system.
Introduction:

Computer simulations are powerful tool and are often used to explain how biomolecular structure and dynamics give rise to biological function. Conventional simulation methods employ explicit representation of solvent environment which can provide a high level of realism but usually at substantial computation costs. Alternatives are implicit solvation models which have been increasingly popular in the simulation of biological macromolecules in order to be able to simulate large system size and longer timescales [168]. Implicit solvation models rely on the assumption that ensemble averages of instantaneous interactions between a solute and solvent molecules may be approximated through mean field formalism [169, 170]. Explicit solvent molecules can then be omitted from the system, thereby reducing the computational cost because of absence of solvent relaxation. Implicit solvent models address the thermodynamic aspects of solvation but neglect hydrodynamic effect that become relevant in the simulation of kinetic processes which have dynamic implications and not only thermodynamic one. Particularly important are stochastic collisions with solvent molecules, frictional forces and viscosity of solvent which directly impact dynamics rates and magnitude of conformational fluctuations. Lacking both these factors results in significant speed up of the sampling process [82]. The entropic or hydrophobic effects arising from solute imposed constraints on organization of solvent molecule around protein surface are not modeled in implicit solvent but can only be approximated through models of non-polar effects like cavity model [27]. The difference in sampling speed and different formulations of both solvents affect conformational sampling in their own way while not
affecting thermodynamic properties being calculated. Conformational sampling can be done at faster rate with implicit solvent like generalized Born with simple switching (GBSW) [23] than explicit solvent. Correct conformational sampling is a key to precise calculation of the thermodynamic properties of protein through computational simulations and timescale over which conformational changes are observed is important for calculating kinetic properties. If the conformational sampling obtained by either of the solvent model is inaccurate then the computed thermodynamic and kinetic properties are not truthful.

The internal motions of protein are essential for their function. Thus an understanding of protein dynamics is of fundamental importance in biology [171]. Protein dynamics is determined by protein energy surface which is a function that describes how the energy of protein varies with the structural changes. For protein configurations similar in structure to the native estate, the energy surface is known to have multiple minima (subsets) and the protein motions at ambient temperature have been shown to involve both harmonic displacements within the minima and across the barriers between them. These conformational subsets have different functional properties [172, 173].

Solvent viscosity has a large impact on these conformational changes and functioning of proteins. For studying effect of solvent viscosity on protein dynamics molecular dynamics simulations has an advantage over experiments that particular solvent properties of interest can be changed without other properties of the solvent being affected. This is particularly true of solvent viscosity. In experiments the viscosity is
usually changed through addition of co-solute, although their molecular size influences the protein dynamics at a given viscosity as shown by Yedgar et al [174]. A change in composition of solvent also implies modification of protein – solvent interactions. In molecular dynamics simulations one can obtain change in solvent viscosity by changing solvent mass, which does not affect other thermodynamic properties of solvent and it does not modify thermodynamic of protein – solvent interactions. Change in solvent viscosity should not affect the equilibrium properties of protein although the dynamic properties of protein will change with change in viscosity of solvent.

Stokes law as shown by,

\[ \eta = \frac{m\gamma}{3\pi d} \]  \hspace{1cm} (4.1)

expresses the shear viscosity \( \eta \) of liquid in terms of the mass \( m \), the friction coefficient \( \gamma \) and the diameter \( d \) of the particles of the liquid. The friction coefficient \( \gamma \) is related to the diffusion coefficient \( D \) of particles through Einstein’s expression

\[ D = \frac{k_B T}{m\gamma} \]  \hspace{1cm} (4.2)

where \( k_B \) is Boltzmann’s constant and \( T \) is temperature, both of which are related to the mean square velocity \( \langle v^2 \rangle \) through equipartition

\[ \frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k_B T \]  \hspace{1cm} (4.3)

Combining equation 4.1 and 4.2 one obtains the relation

\[ D\eta = \frac{k_B T}{3\pi d} \]  \hspace{1cm} (4.4)

Above equation characterizes Stokes – Einstein behavior of liquids. This relation implies that for spherical particles of the same size at constant temperature the product of
diffusion coefficient and shear viscosity is constant and is independent of particle mass. It can be interpreted from equations 4.3 and 4.4 that scaling of total mass of liquid particles by factor of $X$ at constant temperature, it is equivalent to a scaling of the time dimension by factor of $X^{1/2}$ as the kinetic energy is related to mass and time by

$$E_k = \frac{m}{2} \left( \frac{d}{dt} r \right)^2$$

(4.5)

So scaling the mass of particles by factor of $X$, scales the time it takes to travel the same distance by factor of $X^{1/2}$ at a constant kinetic energy. The viscosity should then scaled by factor of $X^{1/2}$, the diffusion constant by factor of $X^{-1/2}$ and their product should be independent of $X$.

With the advantage of implicit solvent models, a number of studies have been performed in recent years [32, 175, 176]. The general conclusion from these studies is that it is possible to obtain stable trajectories from implicit solvent simulations that exhibit conformational sampling comparable to explicit solvent simulations [177, 178]. Some discrepancies, in particular with respect to the stability of salt bridges have been reported, but they appear to be resolved at least in part through careful adjustment of the dielectric interface and/or force field reparameterization [179]. A question that has not been fully addressed to date is to what extent sampling rate of the peptide dynamics or kinetics can be reproduced correctly in implicit solvent simulations and to what extent sampling conformational phase space can be accelerated when solvent viscosity is reduced or omitted. Previous studies have found that conformational transitions in the context of protein folding can be predictably carried using Langevine dynamics with specific friction coefficients, to match the same timescale observed in experiments [35].
Alternatively due to lack of friction and zero effective solvent viscosity, pure implicit solvent simulations can be advantageous in producing similar and thermodynamically accurate dynamics of proteins kinetic processes over much smaller timescale than observed with explicit solvent or in experiments. But this phenomenon needs to be quantified.

In this study we examine effect of solvent viscosity on the rate of sampling peptide dynamics. We compare thermodynamically highly accurate generalized Born simple Switching (GBSW) implicit model with explicit solvent using TIP3P water model for effective time scale required to sample similar conformational changes in model peptides. Polyproline type II helices (PPII) constructed from all proline residues and polyalanine alpha helix (α-helix) constructed from all alanine residues, and Cold shock protein A, derived from *E. coli* were used as model solutes. In addition we use artificial TIP3P water model denoted TIP3P-100 in which atomic masses of TIP3P water are increased by factor of 100. The results for both, dynamic properties like backbone root mean square fluctuations and atom-atom positional cross correlation function sampled at different timescales indicate that the effect of solvent viscosity on protein dynamics also depends on the cross section of solute under observation.
Methods and Materials:

1 Model structures:

Explicit and implicit solvent molecular dynamics simulations of polyproline II helix (PPII helix), polyalanine alpha helix (α helix) and Cols shock protein A (CSPA pdb id 1MJC [180]) were performed in all atom CHARMM 27 force field [147]. The polyalanine alpha helix was constructed with φ = −60° and φ = −45° corresponding to alpha basin and polyproline II helix was constructed with φ = −75° and φ = 150° corresponding to PP-II basin on Ramachandran plot. Both types of helices were made from 10 residues of proline and alanine residues and were capped with carboxyl and amino terminus at ends. The coordinates of CSPA were taken from protein data bank.

2 Implicit and explicit solvent simulations:

In the explicit solvent simulations, the solute was solvated in a cubic box filled with TIP3P water molecules with solute at its center. The resulting box sizes were 40 Å³ for α helix, 54 Å³ for PPII helix and 60 Å³ for CSPA. Migration of the solute protein outside of the primary solvent box was discouraged during the molecular dynamics simulation by weak (0.5 kcal mol⁻¹) center of mass translational restraint using MMFP module of CHARMM [149] on all alpha carbon atoms of protein. Periodic boundary conditions were applied, and particle-mesh Ewald summation with real space cutoff of 10 Å was used to calculate the electrostatic interactions. The explicit solvent system was simulated under isobaric and isothermal ensemble at 300 K with Nose-Hoover thermostat [152, 153] and at constant pressure of 1atm (101,325Pa) maintained by the Langevin
piston method [154]. A time step of 2fs was employed in conjunction with SHAKE [150] in order to constraint bonds between heavy atoms and hydrogen atoms. A standard minimization protocol was applied with initial steepest descent and adopted-basis Newton – Raphson minimization and was followed by heating over 50 ps to from initial temperature 100 K to the final temperature 300 K. Equilibration period of 50 ps at constant pressure of 1 atm and constant temperature of 300 K was applied before production run of 2 ns with 2fs time step in which trajectory was saved every 2ps. Exact same protocol was used for simulations with TIP3P-100 water molecules except longer heating (500ps) and equilibration (500ps) times that were required for equilibration.

Implicit solvent simulations were performed with the GBSW variant of the generalized Born formalism [23]. The dielectric constant inside the solute cavity was set to 1 and to 80 for surrounding medium. A cutoff of 16 Å was used for truncating electrostatic interactions. A time step of 2fs was employed in conjunction with SHAKE [150] in order to constraint bonds between heavy atoms and hydrogen atoms. Both peptides were energy minimized with initial steepest descent and adopted-basis Newton – Raphson minimization. Heating was performed over 50 ps to from initial temperature 100 K to the final temperature 300 K. It was followed by equilibration period of 50 ps at constant temperature of 300 K with Nose-Hoover thermostat [152, 153]. Production run was performed for 2ns at the end in which trajectory was saved every 2ps. Special set of optimized atomic radii developed by Roux et al. was used for implicit solvation simulations [181]. Other atomic parameters were used available in CHARMM 27 all atom force field for both explicit as well as implicit solvent simulations.
3 Normal mode analysis (NMA):

The VIBRAN module within CHARMM force field was used to determine normal mode frequencies by diagonalization of the force constant matrix. Normal modes were calculated for fully minimized structures of both model peptides. Minimization was performed with steepest descent and adopted basis Newton-Raphson method. The structure, as required for NMA, was conformed to be near major energy minimum when rotational and translation modes were zero. No constraints were applied to the system. Trajectories produced by normal mode vibrational analysis served basis for conforming different modes of vibration and RMSF values from NMA.
Results and discussion:

1 Backbone flexibility:

To study the effect of solvent viscosity on peptide dynamics, RMS fluctuations in peptide backbone $\alpha$ carbon atoms positions from average structure were measured at different time lengths. To calculate RMSF at different time lengths, average structures were calculated considering trajectory only for respective time lengths. When the equilibrium is reached, the positional fluctuations of atoms, as measured by the RMSF should not be affected by the different solvent viscosities. However figure 4.1 and 4.2 shows that at shorter time length peptides are more flexible compared to the average structure in GBSW than in TIP3P and TIP3-100. Not unexpectedly, the peptide conformation changes faster in solvent of low viscosity GBSW than in higher viscosity TIP3 and TIP3-100. This is due to in implicit solvent peptide can occupy larger conformational phase space faster in absence of friction coming from solvent. At equilibrium when peptide has sampled conformational phase space with equal distribution, the effect of viscosity on protein motions should be no longer seen. The timescale it take to reach such state will differ and will be longer for higher viscosity solvent. At longest time scale measures, the magnitude of RMSF from GBSW matches with that from explicit solvent TIP3P. TIP3P-100 shows lower RMSF values even at 2ns. The reason for the highest viscosity solvent TIP3P-100, showing lower RMSF than other two is that the simulation is not fully converged in terms of motions captured at equilibrium.
Normal mode analysis (NMA) is useful harmonic analysis method which can capture long timescale, large amplitude and low frequency motions. The peptide motions and RMSF, calculated from NMA trajectory match well with GBSW and TIP3P for PPII helix. The NMA RMSF for α-helix show similarity with that of GBSW but differ in magnitude. RMSF from TIP3P for α-helix does not match with RMSF from NMA especially in the region of tailing residues near n-terminus.

The RMSF in GBSW and TIP3P are converged at 2ns time. As we measure RMSF at different time scale, it can also be noted that time it takes to show convergence of RMSF depends on the solvent viscosity. For GBSW convergence is seen quickest at 100ps while for TIP3P it is longer at 1000ps. For TIP3P-100 it does not seem to be converged and as stated earlier it may take longer than 2ns to converge in term of magnitude that should match RMSF for GBSW and TIP3P. Implicit GBSW reaches the convergence similar to TIP3P faster by about factor of 10 compared to explicit solvent.
Figure 4.1: Backbone α carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for poly-proline PPII helix in different solvents. Reference structure used is average structure generated over respective time scales. RMSF from normal mode analysis (orange line) are shown only in the plot for GBSW and can be compared to all the three graphs.
Figure 4.2: Backbone α carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for poly-alanine α helix in different solvents. Reference structure used is average structure generated over respective time scales. RMSF from normal mode analysis (orange line) are shown only in the plot for GBSW and can be compared to all the three graphs.
An analysis of 2ns simulations for each peptide studied here in implicit solvent GBSW, in explicit solvent using TIP3P water model and artificial explicit solvent with increased atomic masses of TIP3P water designated TIP3P-100, as expected and required by statistical mechanics principles, showed that the equilibrium properties of the peptide should not depend on the solvent viscosity. Though the time scales, over which a particular solvent shows magnitude of motion in terms of RMSF, are markedly different for each solvent. GBSW with complete lack of viscosity shows RMSF similar at 100ps which is very similar to the RMSF shown at 1000ps by TIP3P solvent. Also the same comparison could be made between TIP3P and TIP3P-100. The RMSF of solute in TIP3P-100 at 2ns is much less in magnitude when compared to RMSF of solute in GBSW and TIP3P at 2ns. The RMSF at 500ps of TIP3P solvent matches the RMSF at 2ns from TIP3P-100 solvent. This difference in sampling the phase space at different timescales can be related through simple equation that relates viscosity of solvent to the timescale that it takes to sample the same phase space at equilibrium. As seen from equations 3 and 5, kinetic energy is related to mass and time. So, scaling the mass by factor of X, scales the time by factor of $X^{1/2}$ at constant kinetic energy. The viscosity should then be scaled by factor of $X^{1/2}$, the diffusion constant by factor of $X^{1/2}$ and their product as given by Stokes – Einstein equation, should be independent of X. It can be shown as following. For the two solvents differing in atomic masses by factor of 100, at temperature equilibrium when equal kinetic energy is attended by both solvents we can write,

$$
\therefore \frac{m}{2} \left( \frac{d}{dt} r \right)^2 = \frac{100 \cdot m}{2} \left( \frac{d}{dt_{100}} r \right)^2
$$

(4.6)
From above equations one can expect that solvents which differ in their atomic masses by factor of 100, should have difference in their viscosity by factor of 10 and the molecules of the solvent with higher viscosity will move slower by factor of 10. This situation makes stochastic collisions of solvent molecules with solute slower and less frequent for solvent with higher viscosity, only rate at which solute – solvent interactions occur is affected and solute – solute interactions or interactions within solute are unchanged. This makes the solute dynamics to have impact not by the same factor as with which solvent viscosities and different. This is seen when we compare the RMSF values of solute helices solvated in TIP3P-100 and TIP3P. It is very clearly seen for PPII. The time scale it takes to reach equilibrium in TIP3P-100 is 1000ps while the equilibrium is still not reached at 2000ps in TIP3P-100. In principle, solute solvated in GBSW, which is virtually zero viscosity solvent, completely lack the friction and stochastic collisions from solvent molecules. So the solute is able to explore the phase space much faster. While one cannot relate Stokes-Einstein to the solvent with virtually zero viscosity, we can compare the effect of solvent on peptide dynamics via plotting its dynamic properties sampled at different timescale in explicit solvent and extend the same comparison to implicit solvent as shown in figure 4.1 and 4.2. With our results we observe that PPII is able to sample phase space faster in GBSW by factor of 10 compared to TIP3P and difference between TIP3P and TIP3P-100 to reach equilibrium is by factor of 4. With these results, for simple peptides like isolated helices, one can expect different types of motions observed in implicit solvent to what should be observable on time scale which is over the ten times larger in explicit solvent. Extending the same analogy to TIP3P-100, it should take four
times longer timescale than TIP3P to observe similar motions in helices.

Figure 4.3: Poly-alanine α-helix (on left) and poly-proline type II helix (on right). α-helix is stabilized by array of $i^{th}$ to $i+4^{th}$ residue hydrogen bond (one shown by white dotted line) while PPII helix is stabilized by five member ring side chain structural constraints. The difference in the structure and chain compactness between the two structures is clearly seen in this picture.

2 Differences in helix dynamics as seen by implicit and explicit solvents.

While the general picture for RMSF convergence is true as stated above, implicit and explicit solvent seem to explore little different parts of configuration space in case of
\(\alpha\)-helix as the RMSF for \(\alpha\)-helix from implicit solvent does not match for explicit solvent simulations. In GBSW, \(\alpha\)-helix RMSF are lower and \(\alpha\)-helix seem to stay in configuration very close to its starting structure. While for TIP3P solvent, it deviates more from the starting structure with fraying out last few residues which are close to the n-terminus. Comparing flexibility of PPII and \(\alpha\)-helix, it is noticeable that flexibility of PPII at equilibrium in GBSW as well as TIP3P is very similar and larger in magnitude than \(\alpha\)-helix. Flexibility of \(\alpha\)-helix is different in TIP3P than GBSW and is in general lower in magnitude than PPII helix. This result for \(\alpha\)-helix is almost counter intuitive than expected effect of viscosity of solvent seen on PPII helix. The difference in the behavior of PPII and \(\alpha\)-helix can be attributed to the structure of these two helices.

It can be clearly seen in figure 4.3. PPII helices adopt their conformation primarily because of inherent structure restraints due to its five member side chain envelope conformation [182, 183]. On the other hand \(\alpha\)-helix is stabilized primarily due to \(i^{th}\) to \(i+4^{th}\) residue hydrogen bonds array along the helix giving it net dipole moment along the axis of helix. PPII structure lacks in structural hydrogen bonds completely and the backbone carbonyl and amide dipole orientation are not aligned along the one direction too and does not give a net dipole moment to PPII helix. The backbone atoms of PPII are more exposed to the solvent and PPII gains its stability partially due to the ring side chain constraints and partially due to solvent interactions with backbone [183]. This backbone stabilization of PPII helices makes it more immune to unfolding phenomenon coming from change in environment. Conversely isolated \(\alpha\)-helix, as stabilized solely by backbone hydrogen bonding is more prone towards large changes in conformation and
may undergo unfolding. Though PPII shows flexibility greater than α-helix in terms of backbone RMSF, because of inherent stability of PPII helices, it remains the choice of experimentalists in FRET calibration studies with fluorescent dyes. But due to this flexibility of PPII helices, it has been suggested in literature that for use of PPII as molecular ruler to calibrate FRET efficiency measurements, an accurate determination of probability of its end to end distance is indispensable [184]. In addition, when simulated in GBSW, the hydrogen bonding partners tend to be bonded due to the continuum solvent effect represented by GBSW implicit solvent. This reason is in line with the fact that generalized Born models over stabilizes salt bridges and other electrostatic interactions within solute compared to explicit solvents [179]. TIP3P explicit solvent can compete for hydrogen bonding with hydrogen bond donors and acceptors by explicitly bonding to peptide. Hydrogen bonding within peptide is less frequently observed and shows lesser lifetimes in explicit solvent compared to implicit solvents. This causes α-helix hydrogen bonds to stay in bonded conformation more tightly in GBSW than in TIP3P solvent and makes α-helix more rigid in GBSW than in TIP3P. In this case the effect of viscosity is overcome by the effect of specific model. In case of PPII effect of viscosity is clearly seen on peptide dynamics between GBSW and TIP3P. For both of the peptides though, effect of viscosity on convergence time lengths is very clear when TIP3P and TIP3P-100 are compared since the two models differs strictly only in viscosity and are exactly same in terms of their thermodynamic properties.
Figure 4.4: Ramachandran plots for PPII and α-helix. Dihedral angles are calculated for average for each residue on designated time length of trajectory. The plot on right bottom corner shows a Ramachandran plot with standard regions of \( \phi \) and \( \psi \) dihedral angles for different secondary structural elements for comparison.

3 Evaluation of secondary structure:

Based on the average values of dihedral angles shown in figure 4.4, the relation in timescales between TIP3P and TIP3P-100 is clearer. In TIP3P, some of the regions explored by α-helix are not yet have been sampled by the same solute in TIP3P-100 and
may take longer time. To examine the evolution of secondary structures during simulations, we made Ramachandran plots of the two helices and also calculated RMS fluctuations in \( \phi \) and \( \varphi \) dihedral angles (dRMSF) shown in Appendix A. In general each peptide retains their own secondary structures in all the three solvents as seen from figure 4.4. The biggest difference occurs for \( \alpha \)-helix between GBSW and TIP3P where both solvents tend to explore slightly different conformational space for some of the tailing residues near n-terminus. This is in agreement with the deviation in backbone dRMSF plots in appendix A. In particular 3 residues near to n-terminus of \( \alpha \)-helix forms \( 3_{10} \)-helix or \( \pi \)-helix both of which does not have a significant energy barrier between them and \( \alpha \)-helix structure. These residues also break the hydrogen bonding pattern of the \( \alpha \)-helix structure. Direct exposure to explicit solvent molecules does increase this structural bias. The effect of viscosity when TIP3P-100 is compared to TIP3P is clearly seen too as the secondary structure between the two solvents is similar but shares some differences. The same phenomenon can be seen through calculating RMS fluctuations in backbone dihedral angles (figure 4.4 and Appendix A figure A1 and A2). The last 3 residues of \( \alpha \)-helix deviate immediately from the standard \( \alpha \)-helical dihedral measurements at the start of the simulation and adapt different dihedral angles and stay in that different structure possibly \( 3_{10} \)-helix. They remain in this conformation and keep dihedral RMS similar to other residues.

From figure 4.1 through 4.4 it is clear that as we go from smaller timescale to larger timescales as we analyze of full 2ns run, we start to see that implicit solvent model shows more flexibility than explicit solvent models. The dynamical behavior of peptide
reaches to convergence at very short time (100ps) in GBSW but takes longer time for TIP3P (1000ps) to reach the same convergence. At the same time artificially high viscosity TIP3P-100 solvent has not reach the same equilibrium in terms range of dynamics motion explored by the other two solvents at 2ns. Explicit solvent models appear to be less flexible because of the friction and stochastic collisions with explicitly represented solvent molecules that slow down sampling of dynamics of peptide when compared to implicit solvent simulations. At longer times, the effect of viscosity on sampling is minimized as the motion reach to convergence. But at shorter times, the effect of friction on sampling is seen at maximum.

Another fact emerges that effect of solvent model does not only affect peptide dynamics differently based on viscosity but it also depends on type of peptide or the protein structural element being examined. More rigid systems do not require very long simulations before the explicit and implicit solvent simulations provide similar convergence of conformational fluctuations. As the rigid system show lack of large motions, the solvent effects like viscosity, hydrogen bonding and hydrophobic effects have very slight difference for more rigid system in explicit and implicit solvent. Also due to rigidity, it allows us to sample motions with almost equal rate in both the solvent. In contrast more flexible systems require longer simulations to show the similarity between implicit and explicit solvent flexibility. Flexible systems show more conformational changes. The changing conformations and solvent effects will take more time for complete conformational sampling which will reach to state where it starts showing similar fluctuations in both solvents.
4 Effect of solvent viscosity on protein (Cold shock protein A):

To explore possible effect of viscosity of solvent on type of solute structure used, we simulated a 69 residue β-barrel Cold shock protein A (CSPA) from *E. coli*. The advantage of looking at this protein is it is constituted of different structural elements than helices and that these elements are part of a tertiary structure and not isolated. This allows us a better understanding about the effect of solvent viscosity, when different secondary and tertiary structural elements of proteins are compared among each other. Following figure 4.5 shows RMSF calculated for CSPA in three different solvents.
Figure 4.5: Backbone α carbon root mean square fluctuation (RMSF) in atomic position as measure of backbone flexibility for CSPA in different solvents. RMSF from normal mode analysis (orange line) are shown only in the plot for GBSW and can be compared to all the three graphs.

<table>
<thead>
<tr>
<th>Residue number</th>
<th>1-4</th>
<th>5-13</th>
<th>14-17</th>
<th>18-23</th>
<th>24-29</th>
</tr>
</thead>
<tbody>
<tr>
<td>structure</td>
<td>c-term</td>
<td>β-sheet</td>
<td>loop</td>
<td>β-sheet</td>
<td>loop</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue number</th>
<th>30-33</th>
<th>34-36</th>
<th>37-48</th>
<th>49-56</th>
<th>57-62</th>
<th>63-69</th>
</tr>
</thead>
<tbody>
<tr>
<td>structure</td>
<td>β-sheet</td>
<td>3_10 helix</td>
<td>IDU</td>
<td>β-sheet</td>
<td>loop</td>
<td>β-sheet</td>
</tr>
</tbody>
</table>

Table 4.1: Structural elements in CSPA pdb id:1MJC.

Figure 4.5 and table 4.1 in combination shows effect of solvent in the dynamics of CSPA which is mainly made up of β-sheet, loops and intrinsically disordered unit (IDU) within the protein structure. Close examination of figure 4.5 gives us an idea that as solvent viscosity increased going from GBSW to TIP3P to TIP3P-100, protein dynamics is more damped. Since the protein has different structural elements, as mentioned in table 4.1, not all of them show convergence of RMSF at the same time scale. Loops being more flexible have larger magnitude in RMSF and are not converged at 2ns. RMSF or the movements of different secondary structural elements within protein could be related to their cross section area.
Table 4.2: Diameter (width) of correction area of different secondary structural elements within protein as well as of isolated helices.

<table>
<thead>
<tr>
<th>Structure</th>
<th>α-helix</th>
<th>PPII</th>
<th>β-sheet</th>
<th>loops</th>
<th>IDU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter of cross sec. (Å)</td>
<td>8</td>
<td>5.9</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Above table 4.2 shows cross sectional width of different structural elements within protein as well as isolated helices. From Einstein’s relation of diffusion coefficient,

$$\langle X^2 \rangle = 2Dt = \frac{K \beta T}{6\pi R_h \eta}$$

(4.8)

where $X$ is deviation from reference position, $R_h$ is hydrodynamics radius of particle and $\eta$ is viscosity of surrounding medium. We note that the left hand side of above equation, which is MSD of particles constituting the liquid, is inversely proportional to its hydrodynamic radius. In general this equation suggests that the motion of mobile portion in portion will be inversely proportional to its hydrodynamic radius when submerged liquid media. Relating the same equation to motions involved in proteins, one can expect that the region in protein with larger cross section will be more rigid and will take shorter time so show the convergence on the other hand sections with smaller cross section area will be more flexible and take longest time to show the convergence. The same phenomenon can be seen from figure 4.6.
Comparing data in table 4.2 and following figure 4.6, it shows that in the range of implicit solvent with zero viscosity to explicit solvent with high viscosity as shown in figure 4.6A, all the structural elements which includes isolated helices as well as the β-sheets, loops, and IDU within protein, follow the inverse order of their cross sectional
diameter or width with timescale over which they show convergence, as suggested by diffusion coefficient equation. The order in which secondary structures show convergence turns out to be α-helix, PPII, β-sheet and loops, when compared between GBSW implicit solvent and TIP3P explicit solvent. Intrinsically disordered unit (IDU) in CSPA, which is fairly mobile in GBSW, takes longest time to show convergence but becomes more rigid in TIP3P. The reason could be the stabilization of IDU by specific interaction with water molecules. When we compare TIP3P to TIP3P-100 in figure 4.6B, isolated helices we observe that isolated helices follow expected inverse order of convergence timescale with their cross sectional width. At the same time structures within protein too follow same inverse relation of timescale with cross sectional area. But isolated structure does not follow the order when combined with structures within protein. With this picture now one can relate the structural motions in one particular solvent to expected timescale required for similar motions in other solvent.

5 Effect of solvent viscosity on correlated motions of solute:

Correlated motions of proteins and enzymes have been shown to be important in their functioning and that they are affected by surrounding medium viscosity [185, 186]. The energy surface of protein, like that of any other solutes, is determined by internal potential energy surface and the perturbations due to solvent. The solvent could alter the energy of the protein to create an effective potential energy surface or it could have more direct dynamic effects as a result of collisions between solvent molecules and the protein atoms. We have seen in previous sections how solvent viscosity affects dynamics of
solute peptides and protein. Thus the viscosity dependence of both the effective potential energy surface and the solvent mobility could contribute to transitions. A functionally important characteristic or protein motion is the cross correlation between the atomic fluctuations [187]. This cross correlation analysis demonstrates if atomic fluctuations are correlated (positive correlation), anti-correlated (negative correlation), or uncorrelated (zero correlation). Comparison of atom-atom cross correlation matrix is shown in figure 4.7 for PPI and α-helix and cold shock protein CSP (1MJC) in GBSW implicit solvent, TIP3P, TIP3P-100 and calculated through Normal mode analysis (NMA). The comparison show that rich pattern of correlation and anti-correlation is present in GBSW implicit solvent and fades as solvent viscosity increases. As the viscosity solvent increased all inter-residue ‘communication’ is lost. Based on normal mode analysis, which determines the directions of the proteins with the largest amplitude, the solute in implicit solvent was found to be globally harmonic. Motions along larges principle component or mode observed in implicit solvent are dampened in explicit and high viscosity explicit solvent. This effect is predominantly seen for isolated helices than for protein.
Figure 4.7: Atom-atom correlation function for PPII and α-helix (ALA) and cold shock protein CSP (1MJC). Cross correlation functions are plotted in a matrix for all atoms in the peptide. The x and y axis shows number of atoms while scale on right of each plot shows coloring scheme.
Conclusion:

An analysis of 2ns simulations for isolated helices and different structural motifs within proteins were studied in implicit solvent GBSW, in explicit solvent using TIP3P water model and artificial explicit solvent with increased atomic masses of TIP3P water designated TIP3P-100, is studied here. The results of comparison of backbone RMSF of solute solvated in three solvents indicate that it is not only solvent viscosity and solvent mobility but also type of solute of structural motif under examination, that determines magnitude of protein fluctuation and equivalent timescale at which convergence of fluctuations could be observed. In general high viscosity solvent damp fluctuations in atomic position and solute in implicit solvent with virtually zero viscosity samples similar phase space much faster than explicit solvent and high viscosity explicit solvent. The secondary structure is nearly the same in TIP3P and TIP3P-100 but both shares some differences when compared to implicit solvent. This is due to differences between potential of mean force between them.

One of the main results of this study is it provides comparison between effects of solvent viscosity on different structural motifs. Proteins are made up of different structural elements all of which differ in their structure, stability and contribution to whole protein stability. The rate at which each secondary structure reach equilibrium, accessed in terms of backbone RMSF, depends on its cross section area. Peptide with smaller cross sectional area are more flexible and show highest impact on their potential energy surface as accessed through backbone RMSF at different time scales and atom-atom cross correlation function. On the other hand peptides with larger cross section are
more rigid and are impacted least by solvent viscosity. The order in which secondary structures show convergence turns out to be α-helix, PPII, β-sheet and loops/IDU, when compared between GBSW implicit solvent and TIP3P explicit solvent. All the peptides have inverse relation of their dimension to the magnitude of RMSF.

Correlated motions in proteins and enzymes are shown to be important for their functioning. This study also combines the effect of solvent viscosity on atom positional correlated motions. Again similar to flexibility, high viscosity of solvent dampens correlated motions involved in proteins. For protein or structural element with low cross section and high mobility like PPII, correlated motions are damped the most by solvent viscosity and those structural elements with larger cross section and more rigid the correlated motions are damped to the lowest degree.
Appendix A
Figure A1: RMS fluctuations in backbone dihedral $\phi$ angle (left column of plots) and $\psi$ angle (right column of plots) of PPII helix in GBSW, TIP3P and TIP3P-100 solvents. Since first residue lacks $\phi$ angle, plots in left column extends from 2 to 10 and as the last residue lacks $\psi$ angle, plots in right column extends from 1 to 9.
Figure A2: RMS fluctuations in backbone dihedral $\phi$ angle (left column of plots) and $\psi$ angle (right column of plots) of $\alpha$-helix in GBSW, TIP3P and TIP3P-100 solvents. Since first residue lacks $\phi$ angle, plots in left column extends from 2 to 10 and as the last residue lacks $\psi$ angle, plots in right column extends from 1 to 9.
References:


[184] Doose, S., Neuweiler, H., Barsch, H., Sauer, M. Probing polyproline structure and
dynamics by photoinduced electron transfer provides evidence for deviations from a
regular polyproline type ii helix. *Proceedings of the National Academy of Sciences*

[185] Van Wynsberghe, A. W., Cui, Q. Interpreting correlated motions using normal

[186] Rod, T. H., Radkiewicz, J. L., Brooks, C. L. Correlated motion and the effect of
distal mutations in dihydrofolate reductase. *Proceedings of the National Academy of
Sciences* (2003), 100, 6980-5.

[187] Ichiye, T., Karplus, M. Collective motions in proteins: A covariance analysis of
atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins:*